

Derived Nominals and the Domain of Content

Hagit Borer

Queen Mary University of London

1 Introduction

The starting point of this paper is the investigation of some surprising contrasts between derived nominals with argument structure and event interpretation (Grimshaw's, 1990 Complex Event Nominals), henceforth AS-nominals (for argument-structure nominals) and morpho-phonologically identical derived nominals which do not have event-related properties (henceforth R-nominals). The typology under consideration here, as based on Grimshaw (1990) and much subsequent work is summarized in Table A:

A.	<i>R-nominals</i> ¹	<i>AS-nominals</i> (Complex Event Nominals)
a.	No event-related roles; complements optional	event-related roles; direct arguments obligatory
b.	Event reading not necessary	Event reading.
c.	No agent-oriented modifiers	Agent-oriented modifiers
d.	Subjects are possessives	Subjects are arguments
e.	<i>by</i> phrases are non-arguments; in Spanish, select <i>de</i> ; in Hebrew <i>šel</i>	<i>by</i> phrases are arguments; In Spanish, select <i>por</i> ; in Hebrew <i>ʔal yedey</i>
f.	No implicit argument control	Implicit argument control
g.	(A)telic modification impossible	(A)telic modification possible

While some of the diagnostics are less robust than others, and others are language specific, the typology as a whole remains extremely robust, and claims, as one of its main gains, the confirmed prediction that the mixing and matching of properties across the R-nominals/AS-nominals line should lead to straightforward ungrammaticality, as illustrated in (1). Specifically, in the absence of an internal argument in (1a), *collection* cannot be an AS-nominal, and as a result, the purpose clause is impossible. In (1b) *deliberate* modifies an agent, thereby necessitating an AS-nominal, but *exam* or *journey* can only be R-nominals, leading to ungrammaticality. The telic modifier *in a day* in (1c) requires an AS-nominal, but in the absence of arguments, *destruction* cannot be an AS-nominal, and finally, *disappearance*, presumably derived from the unaccusative *disappear*, cannot be an AS-nominal without its argument realized, and as a consequence, the telic modifier *in two seconds* cannot be licit and ungrammaticality follows, yet again:²

¹ R-nominals in essence following Grimshaw (1990), who assumes that in such nominals (but not in Complex Event Nominals) the external argument is assigned to **R**, a referential index. In the present work, both types of nominals are assumed to be 'referential', in the intended sense. The label is nonetheless preserved as a notational convenience in order to distinguish such nominals from AS-nominals, and should be taken to refer to 'non-AS de-verbal nominals'. Other than their morpho-phonology, identical to that of AS-nominals, R-nominals (so labeled) do not form a coherent distinct group, and there is little to distinguish them in principle, semantically or syntactically, from other derived or underived (non-AS-) nominals (see also section 2.4).

² The most robust of Grimshaw's diagnostics involve argumental *by*-phrases (particularly clear in languages which, unlike English, distinguish it clearly from authorship *by*), implicit argument control, and (a)telicity modification. Note that while the absence of a direct complement is a robust diagnostic for an R-nominal, the presence of a complement does not, in and of itself, exclude an R-nominal construal, and in

1. a. the collection *(of multiple samples) in order to document the spreading of mushrooms
- b. *Mary's deliberate exam/journey
- c. *the destruction in a day
- d. *the disappearance in two seconds

Event denotation, as such, is not restricted to AS-nominals, and is, rather, found in underived nominals as well (cf. 2). Such underived event-denoting nominals, labeled by Grimshaw (op. cit.) *simple event nominals* can be clearly distinguished, however, from event-denoting AS-nominals in at least one important respect. While AS-nominals allow event diagnostics and (a)telicity modification *within* the nominal, the event diagnostics and modification of simple events require embedding the nominal within some additional eventive context (e.g. light verb) and cannot be nominal-internal, giving rise to the ungrammaticality of (3) and to the contrast between (3) and (4):

2. a. The class lasted two hours
- b. The piano recital occurred at dawn
- c. The wedding took place in the morning
3. a. *the class for two hours
- b. *the piano recital in 45 minutes
- c. *the wedding in five minutes
4. a. Mary's teaching of the physics class for two hours
- b. the performance of the piano recital in 45 minutes
- b. Kim's conducting of Kim and Pat's wedding for five hours

Assuming the overall validity as well as the significance of the typology in table A as well as the distinctions between AS-nominals and simple event nominals which emerge from the contrasts in (2)-(4), I will proceed to highlight differences between AS-nominals and R-nominals that cannot be reduced to the specific account given by Grimshaw herself, and which suggest extremely strongly that although one and the same morpho-phonological form, e.g. *collection*, may correspond to the head of both an AS-nominal and an R-nominal, we must assume that those identical morpho-phonological forms have been derived via extremely different routes. Even more importantly, it will emerge that an account for the different properties of such nominals is only available if we assume that complex words are *syntactically* formed, and that the *syntactic* structure that serves as an input to the emergence of an R-nominal is distinct from that which serves as an input to the emergence of an AS-nominal.

In addition to establishing that derived nominals, and by extension complex words in general, must be *syntactically* derived, the account will crucially address central issues that emerge in the context of syntactic approaches to word-formation. Specifically, I will outline here a model for the emergence of non-compositional Content in the context of complex words, following Borer (2009a,b; 2013 a,b) and show how the behavior of de-verbal nominals bears on it and provides evidence for it.³ The account will further provide strong evidence for the claim

fact direct complements are amply present in both non-eventive nominals and in *simple event nominals*, i.e. nominals which happen to denote an event but which are not derived from verbs, such as *event*; *class (of physics)*; *lesson (of physics)*; *concert* and others. See directly below for some discussion of simple event nominals. See fn. 10 for a salient example of the confusion in the literature on some of these issues.

³ Content, rather than e.g. *interpretation* or *meaning* in order to focus attention on the specific way in which substantive vocabulary acquires sense, and which, by assumption, is formally distinct from the assignment of meaning to formal structures or to formal vocabulary. As such, this term presupposes the distinction between *Sinn* and *Bedeutung* as suggested originally in Frege (1892). See (65) and related discussion for a few more comments on the relevant distinction.

advanced in some detail in Borer (2005b), according to which argument structure and arguments, both internal and external, do not emerge as properties of listed items, be they roots or verbs, but rather, emerge from the presence of syntactic structure, and in particular, functional projections which are associated with the verbal extended projection.

2 Two Puzzles

The starting point of the ensuing discussion are two rather puzzling contrasts between AS-nominals and R-nominals which I will now proceed to highlight.

2.1 Puzzle 1: AS-nominals are always compositional (see also Marantz 2001)

The first puzzle concerns the observation that the head of the AS-nominal, i.e. the derived nominal itself, is always compositional, where by *compositional* here I refer to the fact that the Content of the whole can be directly computable from the Content of its parts, and specifically, from the Content of the verbal stem plus whatever function is associated with the nominal suffix. No such restriction holds for the derived nominal when it is an R-nominal (i.e. when it conforms to the diagnostics of R-nominals in table A). To illustrate, consider, e.g. the derived nominal *transformation*. Alongside its fully compositional Content (*the act of transforming*) it also has a non-compositional Content, specifically associated with a grammatical operation within Generative Grammar. That Content cannot be compositionally derived, because the verb *transform* does not have the relevant technical Content. As (5a) clearly illustrates, in its non-compositional, technical sense, *transformation* may not take arguments. No such restriction holds for the compositional sense, as (5b) illustrates (capital italics for non-compositional Content):

5. a. *the *TRANSFORMATION* of the structure by the linguist
(cannot be in reference to *the linguist's having performed a grammatical transformation on the structure*)
- b. the transformation of the field by the linguist

Similar contrasts can be shown to hold for, e.g. *constitution*, *government* and *civilization* among many others:

6. a. *the *CONSTITUTION* of the US by the Founding Fathers (in Philadelphia)
- b. the constitution of the committee by the government (in Philadelphia)
7. a. *the *GOVERNMENT* of the US by Obama
(cannot mean: *the government of the US as put together by Obama*)
- b. the government/governance of the US by the elected president (=act of governing)
8. a. *the *CIVILIZATION* of Europe by Greek and Roman influence
(cannot mean: *the emergence of European civilization as result of Greek and Roman influence*)
- b. the civilization of Europe by Greek and Roman influence
(**can** mean: the civilizing effect that Greek and Roman influence has had on Europe)

–*ing*, when occurring as a nominal suffix, displays the very same properties, and hence the contrast between *READING*, not compositional relative to the Content *READ*, and the virtually

The specific evidence adduced in Borer (2009a) and (2013a) for the model of Content to be outlined here is from the domain of genitive constructions and compounds. While it provides evidence for the extension of the system beyond the domain of de-verbal nominals, a fuller review is clearly outside the scope of this article.

synonymous *interpretation*, compositional relative to the Content of *interpret*, *INTERPRET*, as illustrated by the contrast between (10a) and (10c):⁴

9. a. *the *READING* of the world {*by Aristotle/*for several decades}
- b. the reading of the novel by the students

Transformation, *government* and *reading* are, of course, ambiguous as R-nominals, allowing both compositional and non-compositional construals:

10. a. the transformation was incomplete
 → the change was incomplete (TRANSFORM+ation)
 → the grammatical operation was incomplete (TRANSFORMATION)
- b. Corporate government (=governance) is a tricky affair (GOVERN+ment)
 The government resigned (GOVERNMENT)
- c. the (poetry) reading started at 9:00pm (READ+ing)
 that reading was discredited a long time ago (READING)

It therefore emerges that the generalizations in (11) hold:

11. a. Non-compositional Content → no eventuality interpretation
- b. Eventuality interpretation → compositional Content

While (11) may seem intuitive, in actuality it runs contrary to just about all common theoretical approaches to the formation of complex words and to lexical listedness. Clearly, in any account, and quite independently of lexical or syntactic treatment, the non-compositional Content of *TRANSFORMATION*, *CONSTITUTION*, and *READING* must be listed. What is puzzling, however, is why such listedness should correspond to the inability to take event arguments or event modification. Conversely put, why should the ability to take event arguments and event modification correspond to the absence of listing? Not only is the correlation puzzling, it is incompatible with the logic of lexical listing, as usually deployed in grammatical theories. *Positive* relationship between listedness and argument selection is the cornerstone of lexicalist approaches. E.g. the distinction between the argumental arrays of *give* and *donate*, or *load* and *fill* is *lexically* specified as part of the specific properties of these listed entries. While for the relevant models this might potentially be derived from some lexical semantic differences, there seems to be little reason to assume that non-compositional, listed, derived nominals have, in principle, lexical semantics that is incompatible with event interpretation. To the contrary, e.g. *TRANSFORMATION* and *READING* do have a *simple event* construal, as in (12), and when embedded under light verbs, allow an event interpretation, event modification and event arguments (13):

12. a. [the] Passive *TRANSFORMATION* takes place in two different ways (Google)
- b. The route by which this *READING* emerged is explicitly charged with cultural meaning, and marks a historical moment. (Google)
13. a. The linguist performed a *TRANSFORMATION* on the structure
- b. This *READING*, made by groups with interest in genocide or an equally dominant schema, is in fact in opposition to the suggestion he makes... (Google)

Important lexicalist accounts developed in the 70's and 80's subscribe to the view that derived nominals do not have arguments, as such, and rather, the morphological rule which derives nominals from verbs involves, among other factors, the inheritance of the arguments of the V-stem by the emerging derived nominal (cf. Williams, 1987 and Roeper, 1987). It is worthwhile noting that such approaches would not fare very well in attempting to account for

⁴ By implication, this means that there *are* R-nominals with *-ing*, contrary to Grimshaw (1990) and Marantz (2001). See Borer (2013b) for a detailed discussion.

(11). Specifically, why should lexical inheritance from V be blocked if the Content of the verb is not preserved? All the more so as at least for *TRANSFORMATION* and *READING*, as in (13), the arguments licensed through the light verb are semantically the very same that would occur in compositional AS-nominals (e.g. *agent/patient; originator/undergoer* etc.). The question here, note, is not *whether* inheritance from V would need to be blocked for the relevant cases, for clearly it *is* blocked or (5)-(9) would be licit. Rather the question, within lexical inheritance-based accounts, is *how* to block such inheritance in non-compositional cases, while allowing it in compositional ones, and while holding on to the fact that all occurrences of *transformation* are derived from the existence of some fixed stem, *transform*, which merges with a fixed affix, *-ation*, and which, across the board, gives rise to an identical morpho-phonological output.

The analysis of derived nominals put forth in Chomsky (1970) would likewise face a serious problem here. For Chomsky (op. cit.), there would be a categorially unspecified lexical item, call it *TRANSFORM*, which would acquire its categorial status from its insertion contexts, and where subcategorization would hold constant across its instantiations, in conformity with X'-theory.⁵ That the technical non-compositional Content *TRANSFORMATION* cannot share these arguments can only be stated if *TRANSFORMATION* is altogether banished from the (underspecified) *TRANSFORM* entry, and is effectively no longer assumed to have any relationship with *TRANSFORM*. But if non-compositional derived nominals are now assumed to have an entirely distinct listing from that of their compositional twins, then non-compositional derived nominals, as a matter of principle, can no longer serve as an argument for banning derived nominals from the syntax any more than the distinct properties of [*Nruin*] and *destroy* can serve as an argument for the non-syntactic derivation of compositional *destruction*.

Matters do not improve when we consider the analyses of derived nominals in Marantz, (1997, 2001); Harley, (2009), and Embick, (2010). By assumption, in these accounts roots such as $\sqrt{\text{TRANSFORM}}$ may take an internal argument and *transformation*, in both its compositional and non-compositional instantiations, (may) have either structure in (14a,b), with *n* spelling out as *ation*, and with the (optional) internal argument of the root instantiated either as in (14a) (Embick, 2010) or as in (14b) (Harley, 2009). Either way, it is difficult to see how to exclude either of these configurations just in case $\sqrt{\text{TRANSFORM}}+n$ gives rise to a non-compositional Content:⁶

⁵ But allowing an optional argument for the N instantiation because of categorial properties that distinguish N from V. Note, however that the optionality of the argument cannot account for our puzzle, because it would fail to distinguish between the obligatoriness of the complement in AS-nominals, and its actual exclusion in non-compositional cases.

⁶The specific execution in (14a) is based on the assumption that while the root selects its internal argument, it is not realized as a sister of the root, but rather as a sister of the categorial node (effectively, a specifier of sorts). The structure is particularly suitable for capturing, syntactically, the claim made in Chomsky (1970), according to which arguments are optional in the context of N but not in the context of V. It is important to note, nonetheless, that proponents of the structure in (14a) clearly do subscribe to the view that the root, rather than, say, *v*, is implicated in the assignment of the 'internal' argument, for that is the only way, within such approaches, to derive the well-known contrast between (ib) and (iib). Specifically, Marantz (1997) as well as Harley (2009) and Embick (2010) subscribe to the view that *-ation* nominals are root nominalization, and hence by assumption exclusive of the *external* argument, but allow the nominalization of the root and its selection domain. On the other hand, *-ing* nominalizations crucially scope over *v*, thereby allowing an external argument and rendering (iib) licit:

- | | | |
|--------|--|--|
| i. a. | the growth of the tomato | [intransitive only, root nominalization] |
| b. | *(the farmer's) growth of the tomato (by the farmer) | [transitive, illicit as root nominalization] |
| ii. a. | the growing of the tomato | [intransitive] |

A second equally mysterious puzzle concerns the fact that AS-nominals, but not R-nominals, must always embed a real, attested verb. Specifically, note that the derived nominals in (15) cannot occur as AS-nominals:

15. a. *the vision of the mountain by the instructors for three hours this morning
- b. *the cognition about the problem for three hours by the philosophy student
- c. *the petulance of the passenger for thirty minutes in order to gain access to the lounge

16. *transformacya* *construkcyja* *derivacyja* *administracyja* *glorifikacya...* [M. Hebrew]
transformation construction derivation administration glorification...

17. *gam zavodot ha.bayit hen produkcyā, gam hen yiccur* [Hebrew Wikipedia]
 also works the.house are production also they production
[borrowed] [native]
 'domestic labor as well is production (borrowed), it is also production (native)'

18. *histagglut* *evolucyonit* *o* *adaptacya* *hi* *ha.derek* *še-b.a* *organizmim*
adaptation evolutionary or adaptation she the.way that-in-her organisms
[native] [borrowed]
mitmodedim *zim* *qšayim* *be-sbibat* *ha.mixya* [Hebrew Wikipedia]
cope with difficulties in-environment the.living
'Evolutionary adaptation (native) or adaptation (borrowed) is the way in which organisms
cope with difficulties in their living environment'

These very same vocabulary items, established by the contexts in (17)-(18) to be directly synonymous, nonetheless behave very differently when embedded within AS-nominals, with the native form licit, and the borrowed one directly excluded:⁸

- b. (the farmer's) growing of the tomato (by the farmer) [transitive, nominalization above *-ing*]

⁷ -*cya* presumably originally following Russian -*cje*, and with the final vowel becoming -*a* to merge with the feminine nominal ending in Hebrew. Present day borrowing is primarily from English, but with the ending preserved. Crucially, the 'stems', once -*cya* is substracted (e.g. *konstruk*; *glorific*) are not themselves licit vocabulary items, nor do they occur with any other ending.

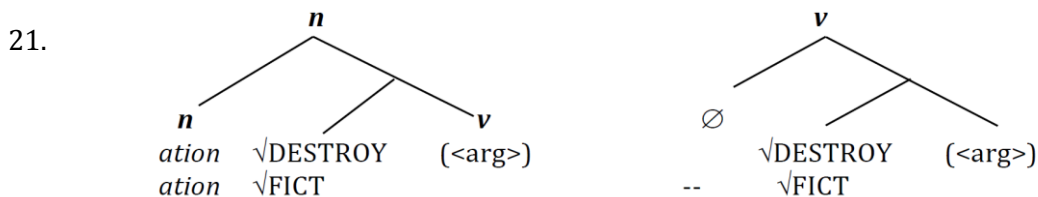
In addition to *-ation*, there are also attested borrowings of *-ure* (pronounced *ura*), although by no means as common. To the best of my knowledge, there are no systematic borrowings of other deverbal derived nominal forms.

⁸ All examples are free nominals rather than construct nominals as borrowed forms, by and large, are uncomfortable as construct heads.

19. a. *ha.yiccur* *šel ha.mexoniot be-ʔesrim šaʔot kedey le.qadem et ha.kalkala*
 the.production of the.cars in-20 hours in-order to advance OM the.economy
 b. **ha.produkcyā* *šel ha.mexoniyot (be-ʔesrim šaʔot) kedey le.kqadem et ha.kalkala*
 the.production of the.cars in 20 hours in-order to advance OM the.economy
 'the production of the cars in 20 hours in order to advance the economy'
20. a. *ha.histagglut* *šel ha.neandertalim le-ʔaqlim eropa tox 20,000 šanim*
 the.adaptation of the.Neanderthals to-climate Europe in 20,000 years
 b. **ha.ʔadaptacyā* *šel ha.neandertalim le-ʔaqlim eropa tox 20,000 šanim*
 the.adaptation of the.Neanderthals to-climate Europe in 20,000 year
 'the adaptation of the Neanderthals to the climate in Europe in 20,000 years'

The need for a 'real' verb as a source for AS-nominals is not, in actuality, a puzzle for a particular kind of lexical execution – the one which associates arguments with listed verbs *exclusively* and which then assumes lexical inheritance (as per discussion in section 2.1). Crucially, for such an execution it must be the case that nouns, as such, never have event arguments or event properties (contra Grimshaw, 1990), and the emergence of any event properties is directly attributable to the presence of a listed verb. Within such an execution [nat] and [vis] would be analytical verbal objects (in the sense of Aronoff, 1976) and hence not listed, thereby excluding arguments. The execution, however, comes at the cost of fundamentally weakening the claim, originally in Chomsky (1970), that derived nominals must be lexically derived and listed, at least for compositional cases.

Once we move away from that particular execution, we find that the need for a 'real' verb in AS-nominals does present a puzzle for the account of derived nominals in Chomsky (1970), where there is no requirement that underspecified entries be fully phonologically realizable as both V and N, and so it is not clear why in the absence of a realizable V the N cannot have a subcategorization frame. It is identically problematic for the reworking of Chomsky (1970) system in e.g. Marantz (1997, 2001) or Harley (2009). Internal arguments, for most of these accounts, are selected by roots, and derived nominals may be derived directly from such roots (see structures in 14 as well as footnote 6). It is not clear how to build into the system the cross-derivational requirement that a root have a licit *v* spellout, in order for its *n* instantiation to be well-formed (cf. 21):



Finally, note that a solution of sorts would be available if one were to assume that *vision* and *fiction* are not derived (i.e. are argument-less roots or underived N) but *destruction* and *adherence* are. Such a move would effectively require a separate, unrelated listing for each member within groupings such as {*native*, *nation*, *natal*}; {*fictitious*, *fiction*}; {*visual*, *vision*, *visible*}, as these cannot be related, morpho-phonologically, either by the existence of a common root such as $\sqrt{\text{NAT}}$ or $\sqrt{\text{FICT}}$, nor by the existence of suffixes such as *-ive*; *-ion*; *-al* or *-ible*. However, once morpho-phonological *relatedness* and *segmentability* are dispensed with as structural criteria, it is no longer clear that morphology would remain a valid subject matter that would be distinct from the investigation of the grammatical relationship between e.g. [_Nclass] and [_Vlearn] which, absent morpo-phonological relatedness and segmentability would be no different from whatever relationship might hold between [_Nstudy] and [_Vstudy] or [_Afictitious] and [_Vfiction].

2.3 A corollary and the phonological faithfulness of roots.

With a solution still pending, it is worthwhile noting a corollary of our puzzles which goes to the heart of the relationship between decomposition and concrete morpho-phonological segmentability. As an illustration, consider the case in (22), or similar pairs such as *die-kill*, *see-show* etc. At the core of analyses that derive e.g. *feed* from [CAUSE EAT] is the explicit or implicit assumption that such decomposition need not be morpho-phonologically faithful, and specifically, that the spellout of EAT need not be consistent across its occurrences. It may be *eat* when e.g. embedded under T, but something altogether different when embedded under CAUSE. Such spellout variability can be directly captured by e.g. late insertion, where EAT would be an abstract lexeme or an abstract root that would spell out late and potentially differently in distinct syntactic environment:

22. a. [EAT] → *eat*
 b. [_V CAUSE EAT] → *feed*

Contrasting with such an approach would be the claim that EAT, whether a lexeme or a root, must be phonologically faithful across its occurrences, or more specifically, that lexemes or roots per force include phonological information that cannot be over-ridden. If that perspective is adopted, then it emerges directly that *feed* cannot be decomposed into [CAUSE EAT], and that whatever it is that relates *feed* and *eat* must be otherwise stated.

Ultimately, the question should be an empirical one, and some evidence should be available to bear on resolving this issue. We note now, and returning to our puzzles, that within an approach that rejects the morpho-phonological faithfulness of roots, or lexemes, the correlation between the need for a ‘real’ verb and the availability of an AS-nominal, mysterious as it already is, is rendered incomprehensible. Specifically, nothing in principle could possibly exclude the derivation in (23c), and the fact that the AS-nominal instantiation of such borrowed forms are systematically missing would become a complete coincidence. To elaborate, in (23c), there is little to exclude a derivation in which the verbal instantiation of some abstract root or lexeme CHANGE would spell out in Hebrew as *šinna*, but when merging with some nominal functor (*n* or C_{N[V]}, see section 3), the output spellout would be either *šinnui* as in (23b) or *transformacya*, as in (23c):

23. a. [_V CHANGE] → *šinna*
 b. [_N[_VCHANGE] *n*] → *šinnui*
 c. [_N[_VCHANGE] *n*] → *transformacya*

It thus emerges that whether or not derivations are phonologically faithful actually has direct syntactic consequences – if one gives up on maintaining such phonological faithfulness, the systematic *syntactic* properties of morpho-phonologically derived nominals cannot be captured, and any hope is lost for characterizing coherently the distinct syntactic properties of *šinnui* vs. *transformacya*, or, for that matter, the absence of AS-nominal instantiation for English forms such as *seminar*, *catharsis* or *activity*, although, in and of themselves, licit as simple event nominals:

24. a. the {*class/seminar/concert/catharsis/activity*} took place at sunset and lasted 90 minutes
 b. the {*class/seminar/concert/catharsis/activity*} (of physics) (*by the instructor) (*to explain the exam) (*for 90 minutes)

In turn, if roots (or lexemes) come with phonological indices, and if AS-nominals must be, indeed, derived from a ‘real’ verb, then both *concert* and Hebrew *transformacya* would be correctly excluded as AS-nominals, as there could exist no licit phonologically faithful verbal realization within them from which they could be derived. Similarly, there could be no such licit phonological verb within *activity*, which although a derived nominal and event-denoting,

As we shall see, once the formulation of our two puzzles is refined, a single answer will emerge for both.

2.4 Who's the culprit?

Note now that R-nominals may be compositional or non-compositional, and that R-nominals may have an otherwise attested verb base or not. In fact, one is hard pressed to come up with syntactic properties of R-nominals that distinguish them from underived nominals, or nominals that are not derived from verbs (or adjectives) altogether. Not so AS-nominals – they must be compositional, and they must be derived from a verb which is otherwise attested as a free verbal constituent. It thus emerges that whatever accounts for our puzzles cannot emerge from the properties of R-nominals, but must find its explanation within the structure and properties that single out AS-nominals.

Importantly, (the head of) R-nominals and (the head of) AS-nominals are morpho-phonologically identical, and overwhelmingly, the very same morpho-phonological derivative is ambiguous between the two. The source of the distinct properties of AS-nominals and R-nominals cannot, then, be morpho-phonological, nor can it involve whatever compositional process puts together, e.g. *TRANSFORM* and *-ation*, given the fact that in both R-nominals and AS-nominals the result can be compositional. Clearly, the source of the difference cannot reside with the base, be it a root or a verb, as it is identical in R-nominals and AS-nominals. Rather, any hope for accounting for the distinct properties of AS-nominals must appeal to the distinct syntactic properties of the larger nominal constituent within which the derived nominal is embedded. Relative to the properties in Table A, and continuing to assume that the relevant properties which single out AS-nominals concern argument structure (including an event argument), the informal statement in (26) thus seems warranted:

26. Event/Argument Structure keeps the AS-nominal compositional

But if the (loose) proposed entailment in (26) is at all on the right track, then it also must be the case that the entailments in (27) hold, rather contrary to much influential work within Generative Grammar for the past three decades:

-
- ii. a. the metamorphosis of the house (*by Mary) (*in order to erase the traces of her ex-lover)
 - b. the scrutiny of dubious looking tax forms (??by the IRS) (*in order to uncover tax evaders) (*for two months)
 - c. the culture of new forms of bacteria (*by my assistant) (*in seven days/*for seven days) (*in order to find an antidote)
 - d. the (official) anathema by the church (*for many centuries) of those taking part in satanic rituals (*in order to protect itself from their challenge)
 - f. the constant mischief by the boy (*for two hours/*in two hours) (*in order to get attention)
 - g. the ongoing treason (*by Quisling) (*in order to support Nazi Germany)
 - j. the homicide of AltaVista and AllTheWeb (*by Yahoo) (*in order to increase the value of its shares)

Impulse, in particular, is of interest because although it may not occur as a stand-alone verb, it does serve as a verbal base for affixation, e.g. as in *impulsive* or *impulsion*, thereby raising the possibility that on a par with *aggress*, this is a case of a real verb, albeit one that is only licit in bound contexts (and note in this context that *aggression* **can** be an AS-nominal). *Recourse*, at present, remains an unexplained counter-example to the claims made here and in Borer (2013b).

Finally, *changeover*, in (i.e) is derived from a verb, making it altogether irrelevant in this context. We note, however, that the occurrence of *changeover* (and *change*) as an AS-nominal is a counter-example to the claim advanced in Grimshaw (1990) according to which zero-derived nominals are never AS-nominals, as well as to the claim in Borer (2013b), according to which English lacks (productive) zero-nominalizers. See Borer (2013b) for fuller discussion.

27. Event/Argument Structure *cannot* emerge from an embedded V
Event/Argument structure *cannot* emerge from an embedded root

More specifically, both embedded verbs and embedded roots behave, morpho-phonologically, identically in AS-nominals and in R-nominals, and therefore cannot be the source of the differences between them, nor can the difference emerge from the properties of the nominal suffix, likewise potentially identical between AS-nominals and R-nominals. The sole difference between them, then, is precisely the set of properties in Table A, i.e. the presence of event/argument structure for AS-nominals, but not for R-nominals. This, in turn, strongly suggests that the distinguishing properties cannot be related to the head of the derived nominal, as such, but must be phrasal in nature. In turn, and insofar as the relevant phrasal properties *cannot* emerge from the properties of either the verb or the suffix, it follows that event/argument structure cannot be an entailment from the lexical semantics of heads, but rather, must emerge as a result of independent structure. Assuming such independent structure to be implicated in licensing syntactic constituents, and hence itself syntactic, the distinct properties of AS-nominals and R-nominals strongly support a top-down, constructionist approach to event structure, rather than bottom up, projectionist system (to use the terminology of Levin and Rappaport-Hovav 1998 and Rappaport-Hovav and Levin, 2005).

2.5 Road signs

In the remainder of this paper, I will focus on offering a solution to the two puzzles thus far presented. Once such a solution has been offered, I will proceed to investigate its theoretical consequences for syntax, for word formation, for argument structure, and for the construction of non-compositional Content.

In section 3, I introduce some notations and assumptions concerning categorial labels and formal vocabulary as based on the model developed in some details in Borer (2005a,b) and Borer (2013b). Crucially, I put forth a view of roots as pure phonological indices, effectively phonological faithfulness devices, which are, in themselves, devoid of any Content and of any syntactic or semantic properties. Section 4 puts forth a schematic syntax for derived nominals, both of the AS- and the R-type. At the end of section 4, our two puzzles are re-cast in terms of the formal mechanisms and the structures proposed.

Section 5 turns to the investigation of how non-compositional Content can be delimited in a syntax-based word formation, clearly a crucial issue in addressing puzzle 1. Assuming, with Arad (2003) and others, that Content must be delimited by a well-defined syntactic domain, I proceed to propose following Borer (2009a,b; 2013a), and contra the claim made in Arad (2003) and adopted by much subsequent research, that the domain of Content is not delimited by the merger of the first categorial label (e.g. *n*, *v*, *a*), but rather by the merger of the lowest segment of an extended projection. Evidence is reviewed both against the claim originally in Arad (2003) for delimitation of Content by categorial merger, and for the delimitation of Content by segments of extended projections. Importantly, much is made of the fact that conventional, listed Content, as such, is neither necessary nor sufficient to give rise to grammatical well-formedness, and that phonological forms which are devoid of encyclopedic Content nonetheless can be shown to obey very clear domain restriction on where, and how, they may acquire such Content.

In section 6 I turn to the resolution of puzzles 1 and 2. Specifically, once the syntactic domain of Content is in place, the obligatory compositionality of AS-nominals, outlined in puzzle 1, will be shown to follow directly from their syntactic structure, as does the possible non-compositionality of R-nominals. As for puzzle 2, it will be shown to vanish altogether, reducing, rather, to puzzle 1. The resolution of both puzzles will be shown to be crucially contingent not only on the correct delimitation of the syntactic domain of Content, but also on assuming that all

direct event arguments are licensed structurally, and that such licensing is oblivious to the properties of both the root and the (lowest) verbal instantiation containing it.

The resolution of puzzles 1 and 2, nonetheless, will be shown to give rise to a third puzzle which concerns the relations between the phonological realization of derivational morpheme and the emergence of non-compositional Content. Section 7 outlines this additional puzzle and proposes a solution which suggests, rather strongly, that there is no unified Intentional-Conceptual interface. While formal semantics takes as its input syntactic structure and properties of grammatical functors, but is oblivious to spellout, the domain of Content, similarly sensitive to syntactic structure, is oblivious to the formal properties of functors, but is highly tuned to their phonological realization.

Section 8 offers a summary and a conclusion.

3 A brief note on assumptions and notation

Within the XS-Model (cf. Borer, 2005a,b; Borer, 2013b), vocabulary, broadly construed, comes in two varieties: roots (*listemes*, in Borer, 2005a,b) and functors. Roots, as such, have no formal (non-phonological) properties. Functors, in turn, are *rigid designators*, in the sense that they name syntactic or semantic functions that are fixed across all their instantiations.¹¹ Functors, in turn, come in two rather distinct formal varieties: Categorical Functors and Semantic Functors.

3.1 Categorical functors (C-functors)

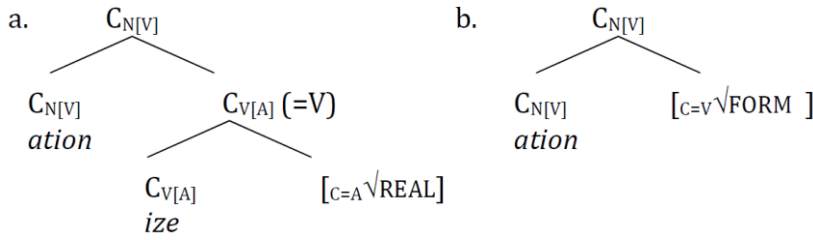
Suppose we view any syntactic structure as a finite ‘space’ which is partitioned into constituents (primarily) by categorical functors (C-functors). Concretely, a C-functor projects a category, and defines a complement categorial space as class-equivalent to another category. Using the notation $C_{X[Y]}$, we now take C to designate a C-functor which projects X, and which defines its complement as equivalent to Y.

Considering a specific example, $C_{N[V]}$ would be a functor which projects N and which defines its complement categorial space as V-(class-)equivalent ($C=V$). In English $C_{N[V]}$ is phonologically realizable (at the very least) as *-ation*; *-ance*; *-ence*; *-ancy*; *-ency*; *-ment*; *-al*. $C_{N[V]}$, in turn, merges with some node, call it M. M may either be, itself, a projection of some C-functor, and hence already categorial, as in (28a) or, alternatively, it may be a category-less root, as in (28b). Either way, M is rendered V-equivalent by being the complement of $C_{N[V]}$. If M itself is already $C_{V[X]}$, as in (28a), for it to be V-equivalent is trivial. If, however, M is a category-less root, being V-equivalent amount to the constituent as a whole having the distribution of V:¹²

¹¹ The term *rigid designator* as used here is thus extended from its original introduction in Kripke (1980), where it is restricted to reference. See, in particular, Gajewsky 2010 on the relationship between (semantic) rigid designation and functional structure.

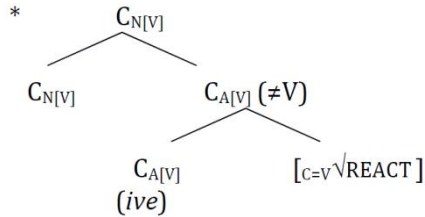
¹² Clearly, movement is required in order for the correct linearization to emerge in (28) and similar cases. For some representations depicting such movement, see section 4 below. For a discussion of the movement assumed here, see Borer (2013b).

28.



(29) is ungrammatical because $C_{A[V]}$, an instance of A, is defined as V-equivalent by $C_{N[V]}$:

29.



C-functors may, but need not, name semantic functions as well. E.g. the C-functor which spells as *-able* in English has (rigidly designating) semantics (and is therefore notated as $ABLE_{A[V]}$).¹³

3.2 Semantic functors (S-functors)

Alongside C-functors, we have semantic functors (S-functors), primarily implicated in the valuing of functional nodes within extended projections. An extended projection is viewed here as a set of segments (ExP-segments) that are valued by S-functors (details aside) and which *collectively* define a particular categorial complement space. By way of an example, $D \in \{Ex[N]\}$, where D is a semantically valued ExP-segment (e.g. by the S-functor THE), and where $\{Ex[N]\}$ is the set of ExP-segments that *collectively* define their complement as N-equivalent. Other members of $\{Ex[N]\}$ would be # (or NUM) and CL (for classifier), and hence $\# \in \{Ex[N]\}$; $CL \in \{Ex[N]\}$. By the same logic, $T \in \{Ex[V]\}$ (valued e.g. by PST) and $ASP \in \{Ex[V]\}$. From our perspective what is of significance is that the categorially formal equivalent of C-functors (i.e. $C_{X[Y]}$) is neither D and T nor THE and PST, but rather the extended projection as a set (i.e. $\{Ex[N]\}, \{Ex[V]\}$) (see Borer, 2013b for a fuller exposition.)

Extended projections, finally, consist not only of the functional sequence, but also of what I shall refer to as a C-core – the domain of C-labeled nodes dominated by the functional sequence. Crucially, the only obligatory part of the C-core is the (C-equivalent) root. In an expression such as (30), however, each of the underlined domains is a C-core, with the highest instantiation of C, in this case $C_{V[A]}$ realized as *-ize*, being the *maximal* C-core. The relevant definitions are in (31):

30. [V] [A] [N] [C=V√TRANSFORM] ation] al] ize]

¹³ We note in this context that the notion of equivalence classes is fundamentally built into all structuralist accounts of syntax, and is at the core of most constituent structure tests that are based on distribution. To wit, to the extent that we typically conclude from the fact that a particular constituent can occur after *the* that it is nominal, or from the fact that it can occur in the subject position that it is (probably) a DP, we are effectively constructing equivalence classes that use syntactic ‘space’ to test for the class equivalence that we might then label N or DP. That in e.g. (29) $C_{A[V]}$ cannot be V-equivalent is thus neither less nor more controversial than claiming that what can occur in the context of T or what can coordinate with V is V, and what cannot occur in the context of D is (probably) not an N.

31. *C-core*_{def}:

- A. α is a C-core iff α is C-equivalent and there is a β such that β is contained in α and β is a root, and for all x , α dominates x and x dominates β , x is C-equivalent.
- B. α is maximal iff there is no γ such that γ is C-equivalent and γ immediately dominates α . (Trivially, recall, all instances of C are C-equivalent.)

3.3 Roots

Roots are pure phonological indices, where a phonological index is in reference to a packet of phonological information which includes the set of all possible (faithful) phonological realizations of the root (allomorphs) as well as the phonological realization of C-functors merging with that root. Thus $\sqrt{\text{GOVERN}}$ would determine the realization of $C_{N[V]}$ as *(govern)ment* or *(govern)ance*, but exclude **governal*; $\sqrt{\text{DEPORT}}$ would allow *-ment* and *-ation* but exclude e.g. *(deport)ance* etc. Crucially, 'faithful' here is meant to be articulated in a manner that would still allow e.g. *catch* and *caught* to realize the same index, but not, say, *eat* and *feed*.¹⁴

Importantly, roots *do not* have Content nor do they have any syntactic properties (e.g. no category, no arguments) or semantic properties. Content, rather, is primarily conceptual. While clearly it can be expressed through natural language, I will assume that the computational system is neither responsible for the emergence of Content, nor is grammar impacted by it in any way. The interpretation which emerges from functors, and in particular from S-functors, is formal and distinct from Content, and neither impacts Content nor is impacted by it in any way.

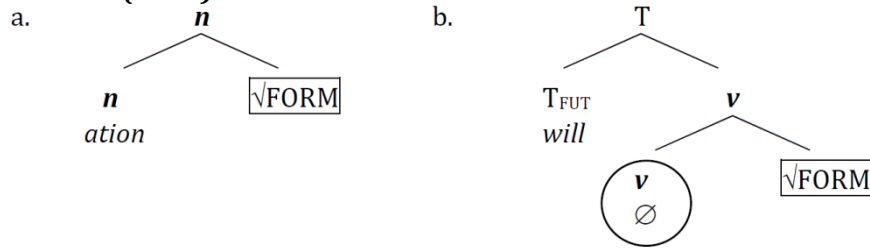
Given the view of syntactic structure as a divided categorial space, non-categorial constituents are never an option. Thus although roots, as such, do not have a category, once embedded within any structure at all, they become C-equivalent. As such, this system differs from the categorial system put forth e.g. in Marantz (2000, 2001) and subsequent, where roots, as such, are never categorized, and where categorial structures are always branching and must involve, in addition to the root, a dedicated categorizer, frequently abstract (i.e. zero-realized *n* or *v*). A specific illustration of the relevant distinct executions is in (32)-(32), with T as a place holder for any member of the verbal extended projection. Note specifically that in (32) the roots (in boxes) remain category-less, and that an additional head, phonologically realized as zero, is necessary to verbalize (or nominalize) $\sqrt{\text{FORM}}$. As a consequence, the verb *form* is always structurally complex. In contrast, in (32), $\sqrt{\text{FORM}}$ is rendered V-equivalent by its syntactic context, that context being either a C-functor or any member of the verbal extended projection, no additional abstract categorial functor is necessary, and the verb *form* does not need to be structurally complex:¹⁵

¹⁴ I am setting aside here as largely irrelevant matters concerning the specific organization of phonological information pertinent to allomorphs, or issues concerning the existence, or lack thereof, of paradigms within such phonological representations.

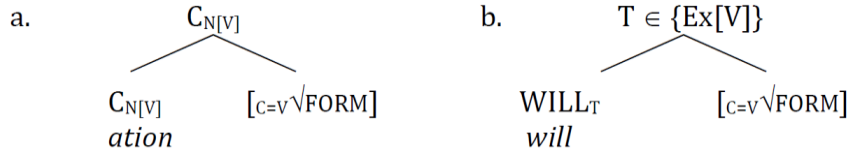
On the issue of faithfulness as relevant to phonological information associated with roots, see Halle and Marantz (1993) as well as Embick and Halle (2005). Note that any such notion of faithfulness would, perforce, bar *go* and *went* from being allomorphs of the same root, a point already touched upon in section 2.3.

¹⁵ The reader should bear in mind that (32b), (32b) are idealizations which abstract away from the possible existence of multiple ExP-segments which dominate the root. Thus relative to (32b), T may at times be the first ExP-segment to dominate the root, but more often than not, it would be another member of the verbal extended projection, e.g. nodes implicated in event structure, aspectual nodes, and others, each of which would, in principle, render the root V-equivalent. To illustrate, if the root is immediately dominated by an ExP segment that licenses a quantity object (ASP_Q in the terminology of Borer, 2005b), it would be ASP_Q , a member of $\{\text{Ex}[V]\}$, that would be rendering the root V-equivalent, and

32. **Marantz (2000) i.a.:**



33. **XS-Model:**



By assumption, any licit sound combination in English is potentially a 'root', a contingency made possible precisely by the claim that roots, as such, do not have Content. Nonetheless, 'listed' phonological indices clearly are more packed with information than unlisted ones and with the latter, it seems, placing no morpho-phonological restrictions at all on what may merge with them (cf. 34):¹⁶

34. a. this more/gribe is much nicer than that more/gribe; the more/gribe is acting up again
 b. he mored and outgribed/overmare and outgrabe
 c. some moration/morance/moring/morement you have here! ; morive; morous; moric; morize; morify...
 d. some gribance/gribment/gribal/gribation!; Outgribious; gribize; gribify; outgribable...

4 **A (simplified) syntax for derived nominals**

Let us now assume that the structure of R-nominals is as in (35)-(36):¹⁷

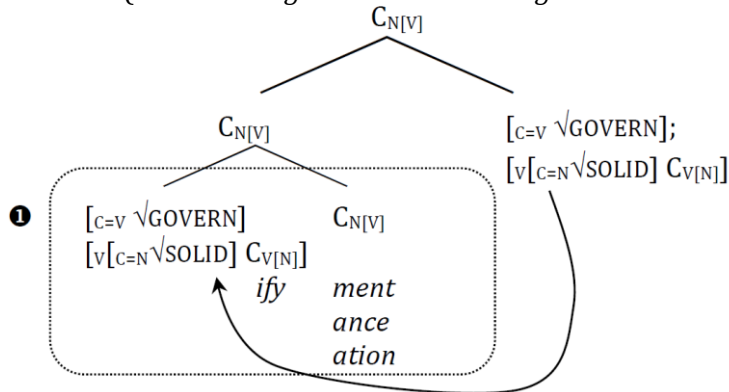
with the structure in (i) resulting for, e.g. *the crystals formed* or *the chemist formed the crystals*:

- i._[ASP-Q] DP ASP_Q [C=V√FORM]
the crystals

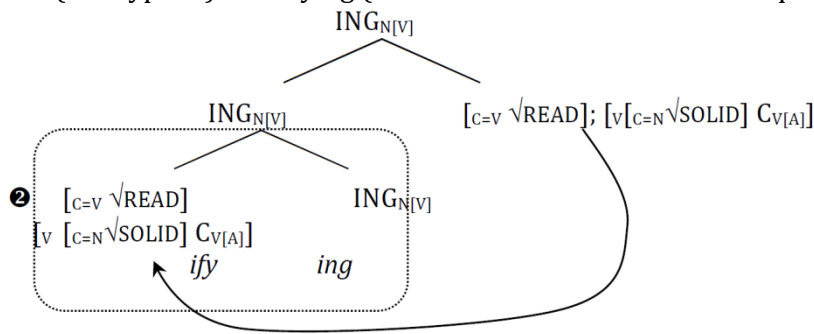
¹⁶ If the freedom of affixation attested in (34) is theoretically significant, it suggests that selection, for 'listed' roots, is a filter, excluding some realizations in a local context, rather than a positive selection. E.g. rather than select *-ance* and *-ment*, $\sqrt{\text{GOVERN}}$ would **exclude** *-ation*, and *-al*. The point is noted here, but not pursued.

¹⁷ Recall that I assume *-ance*, *-ment*, *-ation* and *-al* to be distinct, suppletive exponents of a single nominalizing, deverbal C-functor, itself semantically inert (and hence $C_{N[V]}$). Not so *-ing*, which is the realization of a distinct C-functor, $\text{ING}_{N[V]}$. To be sure, $\text{ING}_{N[V]}$ projects N and merges with V, but as is argued in some detail in Borer (2013), it also contributes semantic content, and specifically brings to the derivation an *Originator* role as well as a non-culminating event interpretation.

35. a. (their) government/governance (is perfect)
 b. (the) solidification (was complete)
 (→Note that *government* but not *governance* can be non-compositional)

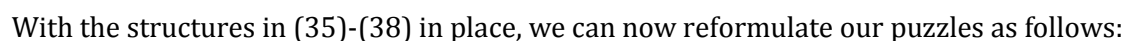
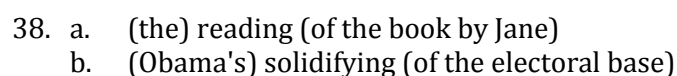


36. a. (that) (political) reading (misses the significance of the Arab Spring)
 b. (this type of) solidifying (occurs whenever iron cools too quickly)



The structure of AS-nominals, however, is rather different. Specifically, and in line with the detailed argumentation in Borer (2005b), I will assume that event structure is licensed through dedicated functional structure. Moreover, the relevant functional nodes here are ExP-segments which are part of the verbal extended projection, i.e., X and Y in (37)-(38) are members of $\{\text{Ex}[V]\}$. Setting aside, for expositional reasons, the specific semantics of the nodes involved or how they are valued, the emerging configuration is as in (37)-(38), with (optional) $Y \in \{\text{Ex}[V]\}$ implicated in licensing a direct internal argument, and with $X \in \{\text{Ex}[V]\}$, by assumption, a node that licenses the event argument as well as, potentially, an external argument:¹⁸

¹⁸ While there is a relatively broad agreement within constructionist approaches on the need for functional syntactic representation for argument/event structure, there is no clear consensus on the specific labeling of the nodes involved, on their overall number, or on their semantics, nor is the claim that internal arguments, specifically, are structurally licensed universally held. For the purposes of the present article, neither the labeling nor the semantics of the relevant nodes is relevant, and this matter is thus set aside with X and Y opted for, instead. What is crucial, however, is that the nodes implicated, whatever they turn out to be and however many of them there may be, are ExP-segments of $\{\text{Ex}[V]\}$ (and thus render the root V-equivalent), and that both internal and external argument are licensed by such ExP-segments without any recourse to root-listed information.

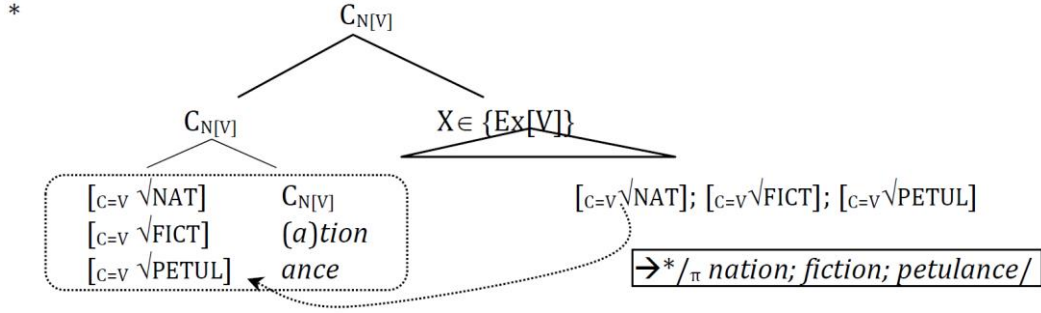


39. I. Why can boxes ❶ and ❷ correspond to non-compositional Content, but not so boxes ❸ and ❹ (puzzle 1)?
- II. Assuming that the structure in (40) is licit, and is thus equivalent to the structure in (35), why is (41), an AS-nominal and the (abbreviated) structural equivalent of (37), excluded? (puzzle 2)

Diagram illustrating the decomposition of the Noun Class (NC) into its constituent parts:

- Root: $C_{N[V]}$
- Branches: $C_{N[V]}$ and $[C=V \text{ NAT}]; [C=V \text{ FICT}]; [C=V \text{ PETUL}]$
- Sub-branches (from $C_{N[V]}$): $[C=V \sqrt{\text{NAT}}]$, $[C=V \sqrt{\text{FICT}}]$, and $[C=V \sqrt{\text{FICT}}]$
- Grouped sub-branches (indicated by a dotted box): $[C=V \sqrt{\text{NAT}}]$, $[C=V \sqrt{\text{FICT}}]$, and $[C=V \sqrt{\text{FICT}}]$ are grouped together, with the label $C_{N[V]} (a)tion\ ance$.
- Resulting words (indicated by an arrow): $\rightarrow nation; fiction; petulance$

41.



Note now that in both R-nominals and AS-nominals, $C_{N[V]}$ or $ING_{N[V]}$ merge with an identical base, and in both, that base is V-equivalent. However, in R-nominals, the base is rendered V-equivalent by virtue of being the complement of the nominal suffix itself, i.e. $C_{N[V]}$ or $ING_{N[V]}$. Not so in AS-nominals, where the base is rendered V-equivalent as a result of being the complement of $\{Ex[V]\}$, a verbal extended projection, and where what merges with $C_{N[V]}$ or $ING_{N[V]}$ is a constituent already previously verbalized. Taking this to be the key difference between the structure of AS-nominals and R-nominals, the following puzzling descriptive generalizations emerge:

42. A. $\{Ex[V]\}$ (or some subset of it) enforces compositionality. Why? (Puzzle 1)
 B. $\{Ex[V]\}$ (or some subset of it) enforces a 'stand-alone' phonologically faithful V. Why? (Puzzle 2)

Interestingly enough, a third puzzle emerges once our structures are in place and the descriptive generalizations explicitly stated. Clearly, *-ment* and *-ance* are no more than alternative spellouts for $C_{N[V]}$ in the context of $[\sqrt{GOVERN}]$. If that were not the case, we would have expected the AS-nominal (37a) to have a different Content, depending on the choice of *government* vs. *governance*. What we find, however, is that the AS-nominals has an identical interpretation, regardless of the choice of *government* or *governance*. Once we turn to R-nominals, the picture becomes more complex. While under a compositional construal the R-nominals *government* and *governance* are synonyms, this need not be the case. Rather, as an R-nominal, and exclusively so, *government*, but not *governance*, may have a non-compositional Content. But if *-ance* and *-ment* are no more than alternative spellouts for $C_{N[V]}$ in the context of $[\sqrt{GOVERN}]$ giving rise to a possible synonymy even for R-nominals, how can the specific non-compositional Content *GOVERNMENT* emerge in the context of *-ment* but not in the context of *-ance*?

5 The syntactic domain of Content

5.1 En-searching and Content matching

Clearly, any attempt at resolving our three puzzles must start with a better understanding of how Content comes to be associated with grammatical constituents, be they terminals or complex structures. By assumption, functors do not have Content but are, rather, grammatical functions. Nor do roots have Content, being, as I suggested, pure phonological indices. Where and how, then, does Content emerge, and how does it come to be associated with strings produced by the computational system?

By way of an answer to these questions, suppose we assume the system developed in Borer (2009a,b; 2013a,b), the essence of which is hereby outlined:

There exists a reservoir of (conceptual and/or potentially conventionalized) atomic, indivisible Content units, call it the *encyclopedia*. Mediating between the encyclopedia and grammatical representations is a search engine (*en-search*), which matches qualified bracketed (at least partially spelled out, as we shall see) constituents with encyclopedic Content units.

II. Non-Compositional Content (=atomic Content) is the output of a *single, successful en-search*, and could be equally associated with *cat*, by assumption a non-branching structure, as it may be associated with *transformation* or *naturalize*. Roots, *phonological indices*, may happen to be co-extensive with a qualifying (partially spelled out) constituent which is matched with Content, but do not, as such, have an independent Content of any sort, nor do they represent a privileged domain for Content matching.

43. a. *slithy, mimsy* b. *sturdy, flimsy* c. *bloody, flakey* d. *bulgy, bossy*

44. $\left[\begin{array}{c} \text{A} \\ \text{C=N} \sqrt{\text{XYZ}} \\ \text{C}_{\text{A[N]}} \\ (s)y \end{array} \right]$

B. <i>BASE (N-equivalent, co-extensive with root)</i>	<i>Content</i>	<i>Derivative</i> [[_{C=N} $\sqrt{\text{XYZ}}$] C _{A[N]}] (s)y	<i>Content</i>
a. <i>slith, mim(s)</i>	No listed Content	<i>slithy, mimsy</i>	No listed Content
b. <i>sturd, flim(s)</i>	No listed Content	<i>sturdy, flimsy</i>	<i>STURDY, FLIMSY</i>
c. <i>bulge, boss</i>	<i>BULGE, BOSS</i>	<i>bulgy, bossy</i>	No listed Content
d. <i>blood, flaked.</i>	<i>BLOOD, FLAKE</i>	<i>bloody, flakey</i>	<i>BLOODY, FLAKEY</i>

The picture in Table B is by no means surprising or new, and illustrates the fact that both base and derivative may or may not have listed Content, and that the Content listed for the base may or may not inform that of the derivative. It is nonetheless worthwhile making explicit at least some of the properties of Content matching as they emerge from Table B.

¹⁹ *Slithy* and *mimsy* are from *Jabberwocky*, a poem in Lewis Carroll's novel *Through the Looking Glass*. The poem is a well-known exercise in constructing text which appears grammatical and meaningful, but in which all Content items are nonce forms. In the original tale, Humpty Dumpty 'explains' the poem to 6-year old Alice, but his explanation consists of matching all these nonce items with Content which is otherwise unattested in any English dictionaries.

BOREER 2013

correspond to any conventionalized listed Content. That, in and of itself, clearly doesn't block their merger, e.g. with $C_{A[N]}$, giving rise, we assume, to the structure $[_{C=N} \sqrt{\text{SLITH}}/\sqrt{\text{MIM}}(S) - C_{A[N]}$ and with $C_{A[N]}$ spelling out as $-(s)y$. The emerging derivatives, *slithy* and *mimsy* are similarly devoid of listed Content, yet again something that doesn't block them from occurring in the otherwise perfectly syntactically well-formed first stanza of *Jabberwocky*. Content matching, then, is clearly optional.

Moving on to *sturd* and *flim(s)*, again, by assumption, N-equivalent constituents co-extensive with $\sqrt{\text{STURD}}$ and $\sqrt{\text{FLIM}}(S)$, they, too, are not matched with any listed Content. However, Content *is* matched with the derivatives *sturdy* and *flimsy*. That Content clearly cannot be compositional for the simple reason that *sturd* and *flim(s)* have no Content. It thus emerges that Content matching may pick a branching constituent and yet match it with atomic Content. Insofar as Content matching is optional to begin with, the fact that *sturd* and *flim(s)* were never matched with Content presents no problem, and Content matching to the larger constituent can proceed directly.

Diametrically opposed to *sturdy* and *flimsy* is the case of *bossy* and *bulgy*. Here listed Content is available for the stems, specifically as *BOSS* and *BULGE*, but arguably not so for the derivative. *Bossy* and *bulgy*, as a result, only have whatever Content emerges from composing the listed Content *BOSS* and *BULGE* with a (Contentless) C-functor, and with full compositionality emerging as a consequence.

Finally, a more complex case is presented by *blood(y)* and *flake(y)*. Here, both compositional and non-compositional Content is available. For the former, the most obvious analysis would be to assume that Content is matched with *blood* and *flake* (*BLOOD* and *FLAKE*), and that the compositional derivative as a whole emerges along the same line of *bulgy* and *bossy*. For the non-compositional Content, (*BLOODY*=*DAMN*; *FLAKEY*=*FLIGHTY*, *SCATTER-BRAINED*) on the other hand, the most obvious analysis would be on a par with *sturdy* and *flimsy*. Specifically, and precisely *because* Content matching is optional, we may assume that no Content was matched with *blood* and *flake*, but subsequently, Content *was* matched with the derivative as a whole.

Three important consequences now emerge. First, we must assume that Content matching is not only optional, but also cyclical. Because Content could, and was, matched with the base, *flakey* and *bossy* could have a compositional Content. Because Content could fail to be matched with the base, but *could* be matched with the derivative as a whole, *sturdy* and *bossy* could be matched with atomic Content, effectively making them non-compositional in the relevant sense.

Another important and by no means trivial consequence emerges when we consider the cases in (45). Note, specifically, that while it was possible for *sturd* and *flim(s)* to be integrated into a Contentful sequence once they merged with $C_{A[N]}$ and thus to give rise to *sturdy* and *flimsy*, no 'rescue' so to speak is available once they are embedded under ExP-segments, be they members of $\{\text{Ex}[N]\}$ or of $\{\text{Ex}[V]\}$. Rather, in (45), *sturd* and *flim(s)* are doomed to permanent Contentlessness. Similarly, of course, *slithy* and *mimsy*, where no rescue is available by embedding them under e.g. the degree phrases as in (45d,e):

45. **Permanently Contentless:**

- a. (the) *sturd(s)*
- b. (will) *flim(s)*
- c. *sturd(ed)*
- d. very *slithy*
- e. *mimsy* enough!

In and of itself, the conclusion that items such as determiners or past tense marking delimit non-compositional Content is yet again hardly new, and corresponds directly to the well-known claim that 'inflection' (broadly construed) doesn't impact Content, but 'derivation' (equally

broadly construed) does. Within a syntactic approach to complex words, however, it indicates that the distinction between functors such as those realized as *the* or *ed*, and functors such as those realized as *-ation* and *-(s)y* must be formal and must be carried into the computational system even when the architecture of words is integrated back into the syntax. Specifically, what emerges is that while the domain of Content matching can be extended to include C-functors, it is restricted by the merger of ExP-segments, or, differently put, the first merging ExP-segment constitutes the upper boundary of en-searching.

Finally, note that e.g. *flakey* is not, in and of itself, ambiguous in any given context. Rather, in any given derivation it is either compositional, i.e. denotes the property of having *flakes*, or non-compositional, i.e. denoting the property of being flighty or scatter-brained. While en-searching may clearly return both Content for *flake*, and thus *FLAKE+y* and for *flakey*, and thus *FLAKEY*, it is clearly only *one* Content in any given derivation – *flakey pastry* is not simultaneously *FLAKEY*, nor is a *flakey guy* shedding *FLAKES* in any sense. It thus emerges that only one Content can be matched with any qualifying string, and that over-riding existing Content is neither possible nor necessary. We note, again, that the uniqueness of Content for any qualifying string, including complex ones, is possible precisely *because* Content matching is optional. Were Content matching obligatory, the matching of the Content *FLAKEY* would have required explicitly over-riding or re-analyzing the Content *FLAKE*, or would have given rise to ambiguity.

Before moving on, it is worthwhile pausing to consider more broadly the logic of argumentation from nonce forms. Importantly, I am not claiming here that Content *cannot* be matched with e.g. *slithy* or *sturd*. To the contrary, it most certainly can, and is an exercise which Humpty Dumpty undertakes with great flourish (cf. fn. 19). The point that I do seek to make, rather, regards any particular speaker and the Content matching possibilities available to him/her given some fixed starting point. By way of a thought experiment, consider the individual Humpty Dumpty (HD), whose *en-search* returns Content for *slith* as well as for *sturd* in a nominal context, say *SLITH* and *STURD* respectively. We now predict that HD will be perfectly comfortable with *slithy* and with *sturdy* having the compositional Content *SLITH-y* and *STURD-y* respectively. Finally, it is also possible that HD has atomic Content units which correspond to *slithy* or *sturdy* as a whole, but which have Content which is distinct from the compositional Content of *SLITH-y*; *STURD-y*, and having, say, the Content *SLIMY* and *STURDY* respectively. What we do not expect, however, is the emergence, for HD, of Content for either *slith* or *sturd* which is distinct from *SLITH* or *STURD* in contexts such as *every sturd*, *three sliths*, *the sturd* or *a slith*.

Suppose, we consider now another individual, say LC (for a hypothetical Lewis Carroll), whose *en-search* component returns Content for neither *slith* nor *sturd*. Under plausible assumptions, most writers of English dictionaries as well as Lewis Carroll belong in that category. What we now predict is as follows: for LC, embedding *slith* or *sturd* under a determiner or under a quantifier will not change matters much, and specifically no Content will emerge for e.g. *the slith*, *these sturds*, *three sliths*; *every sturd* etc. However, we do allow for the possibility that *en-searching*, for LC, will return Content for *slithy* or *sturdy* (as well, of course, as for *slithal*; *sturdal*; *slithify*; *sturdify* etc.). Importantly, such an *en-search* could be successful although the *en-search* based on *slith* and *sturd* has failed. Following that very same logic, whatever Content is matched with *sturdy* and *slithy*, for LC, would necessarily be atomic and non-compositional, for the simple reason that neither *slith* nor *sturd* correspond to Content in LC's encyclopedia. Finally, we also predict that if, for LC, an *en-search* based on e.g. *dump* returns Content, that Content would be one and the same in *the dump* and in *three dumps*, and just in case *dummy* is morpho-phonologically possible, we predict it to allow for the

compositional Content *DUMP+y*, possibly, but not necessarily, alongside some non-compositional Content, say *DUMPY*.

To summarize:²⁰

46. a. Content matching is optional (*sturd*; *slithy*)
- b. Content matching is cyclical. Thus, effectively, Content matching need not be to the maximal C-core (*FLAK+y* as well as *FLAKEY*); but ---
- c. One Content per C-core (*FLAKE* or *FLAKEY* but not both).
- d. The first merging ExP-segment defines an upper boundary on en-searches, i.e., the maximal domain which can serve as an input to a single *en-search*. As a result, and at the string *sturd* may still be integrated into a single successful en-search when embedded within *sturdy*, but not when embedded under e.g. PST and occurring as *sturded* or when embedded under D, and occurring as *the sturd*.

5.2 Non-compositional Content in perspective

Before turning to the resolution of our puzzles, it is worthwhile to digress briefly to consider non-compositional word Content from a broader theoretical perspective. Beginning with Chomsky (1970) and reinforced by much subsequent work, it is typically assumed that non-compositional Content is the hallmark of lexical operations, and that syntactic operations per force give rise to compositionality. We note now that while there is no question that non-compositional Content must be somehow listed, it does not in actuality follow from such listedness that word formation must reside in some component of the grammar, call it the lexicon, which is distinct from syntax. Rather, what follows is that non-compositional, atomic Content may be matched with some domains but not with others. That the distinction between these domains requires distinct grammatical modules with distinct formal properties is neither entailed nor, for that matter, is it obvious. Rather, a domain delimitation can emerge from locality, in and of itself a classical syntactic effect, i.e. in allowing Content matching in some *syntactically* delimited domains, but not in others. What is required, from such a perspective, is identifying the relevant local domain, itself constructed by whatever operations put together phrases (e.g. *Merge*), which interacts with Content matching. The proposal put forth here and summarized in (46) amounts to exactly that – it postulates an independently well-defined local syntactic domain which is accessible to a single en-search, and which is delimited by the occurrence of ExP-segments.

²⁰ An obvious question concerns the existence, in particular, of pluralia tantum with Content that is clearly distinct from that of the non-plural form, such as *premises* (vs. *premise*), or *briefs* (vs. *brief*). The matter is addressed in some detail in Borer (2013a) in the context of a phase-based approach to Content matching. Specifically, I note that within a Bare Phrase Structure approach, an (embedded) constituent cannot be, formally, a maximal projection before it merges with a head. As a consequence, equating a phase with a maximal projection, with any label, is incoherent, and rather, at the very minimum a phase must contain the boxed structure in (i):

i.
$$\begin{bmatrix} \text{C} & \text{C} & \text{B} & \text{X} & \boxed{\begin{bmatrix} \text{B} & \text{A} \end{bmatrix}} \end{bmatrix}$$

Suppose we assume now that B, in (i), is an ExP-segment. The result which emerges is that B, potentially a realized S-functor, *may* be a phase-mate of A, but not so, e.g. C, or, for that matter, X, the specifier of B. The precise statement of the generalization in (46d), then, is that the upper boundary for en-searching is defined by the first phase containing the maximal C-core, or, more precisely, is delimited by the first ExP-*bracket*. If we assume plural markers to realize the ExP-segment which directly merges with the C-core (in contrast with e.g. determiners), the emerging prediction is that plural markers, but no higher functional realizations, would enter en-searching. As is argued in Borer (2013), similar conclusions follow for e.g. Slavic perfective prefixes, by assumption constituting the first ExP-segment to merge with the verbal C-core.

That Content matching (in our terms) must be local in the relevant sense was first proposed, to my knowledge, in Arad (2003) following ideas in Marantz (2000). The local domain proposed, however, differs from that outlined here, and is, rather, assumed to be linked to the first instance of categorization, or within the specific system of execution in Arad (op. cit.), bound by the first occurrence of a categorial node such as *n*, *v*, *a*, etc. As it turns out, however, the domain thus defined is empirically too narrow. An investigation of English attested vocabulary reveals directly that non-compositionality may emerge well past categorization:²¹

47. a.	[_N [_N [_v react] tion] ary]	REACTIONARY	≠	REACT+tion+ary REACTION+ary
b.	[_v [_A [_N natur] al] ize]	NATURALIZE ²²	≠	NATURE+al+ize NATURAL+ize
c.	[_N [_v [_A civil] iz] ation]	CIVILIZATION	≠	CIVIL+ize+ation CIVILIZE+ation
d.	[_N [_A [act] ive] ism/ist]	ACTIST/ISM	≠	ACT+ive+ist/ism ACTIVE+ist/ism;
e.	[_N [_N [_v protect] ion] ism]	PROTECTIONISM	≠	PROTECT+ion+ism PROTECTION+ism
f.	[_A [_N [_v except] ion] al]	EXCEPTIONAL(=EXCELLENT and compare with SPECIAL)	≠	EXCEPT+ion+al EXCEPTION+al
g.	[_{N/A} [_N [_N [_v edit] or] y] al]	EDITORIAL	≠	EDIT+or+y+al EDITOR+y+al EDITORY+al
h.	[_v [_{N/A} [_N [_N [_v edit] or] y] al] ize]	EDITORIALIZE	≠	EDITORIAL+ize
i.	[_N [_A [_v relate[ive] ity]	RELATIVITY	≠	RELATE+ive RELATIVE+ity

We noted already that a C-core may correspond to more than one atomic Content, but never within a single derivation. As a vivid illustration, consider *civilization* in (48) (and assuming for expositional purposes that the root is [_C√CIVIL] rather than [_C√CIV], otherwise attested, e.g. in *civic*). Here, as is rather clear, neither *CIVIL* nor *CIVILIZE* are part of the Content *CIVILIZATION* under its *CULTURE/SOCIETY* reading (e.g., the Mayan Civilization). The point is all the more clearly driven home by the fact that *civilization* may have a Content compositional on both *civil* (i.e. the act of making *CIVIL*) and on *civilize* (the act of *CIVILIZING*, *ENLIGHTENING*):

48.	en-search	Derivation 1	Derivation 2	Derivation 3
i.	[_C √CIVIL]	CIVIL	no listed Content	no listed Content
ii.	[_v [_{C=A} √CIVIL]ize]	CIVIL-ize	CIVILIZE (=ENLIGHTEN)	no listed Content
iii.	[_A [_{C=A} √CIVIL]ize]ation]	CIVIL-ize-ation	CIVILIZE-ation (=ENLIGHTENING)	CIVILIZATION (=SOCIETY/CULTURE)

²¹ The situation is even clearer in a language such as Hebrew, where most (but not all) verbs are already complex and include both the root and a ‘categorizer’ and where, as a consequence, most nominalizations apply to an already categorized constituent. And yet, non-compositional Content for derived nominals is extremely common, and no clear contrast emerges between those which embed derived, complex verbs, and those which do not. See Borer (2013) for detailed illustrations.

²² A. Marantz (p.c.) notes that a “natural citizen” is a technical concept in immigration law, thereby rendering *naturalize* compositional. The compositionality of *naturalize*, however, only follows if we assume that any individual familiar with *naturalize* with the Content *NATURALIZE*, is also familiar with the term *natural citizen* with the relevant Content. Given the fact that the non-compositional Content of *NATURALIZE* is in wide use but the legal concept “natural citizen” is familiar to very few, the rationale here may have some historical merit, but no synchronic force.

That a C-core may only be associated with a single successful en-search within any given derivation receives particular confirmation from the contrast between complex derivatives and compounds. While complex derivatives such as *civilization* or *naturalize* have a single C-core, compounds may have two or more C-cores, and as a result, more than a single atomic Content may be associated with a single derivation. To see that this is the case, consider first the matching of Content with a perfectly compositional compound such as *kitchen towel*, with three C-cores - one co-extensive with $\sqrt{\text{KITCHEN}}$, one co-extensive with $\sqrt{\text{TOWEL}}$, and the third consisting of both:

49. $[_{N1} [_{N2} \sqrt{\text{KITCHEN}}] \sqrt{\text{TOWEL}}]$ (two C-cores plus the maximal one)

<i>(en-search domains boxed)</i>			CONTENT AND COMPOSITION:
i.	en-search 1	$[_{N1} [_{N2} \sqrt{\text{KITCHEN}}] \sqrt{\text{TOWEL}}]$	$\text{KITCHEN} \quad \sqrt{\text{TOWEL}}$
ii.	en-search 2	$[_{N1} [_{N2} \sqrt{\text{KITCHEN}}] \sqrt{\text{TOWEL}}]$	$\text{KITCHEN} \quad \text{TOWEL}$
iii.	en-search 3:	$[_{N1} [_{N2} \sqrt{\text{KITCHEN}}] \sqrt{\text{TOWEL}}]$	no listed Content

Compare now the matching of Content to *kitchen towel* with the matching of Content to e.g. *dragonfly*, as in (50) and to the compound *walkman* as in (51) (note that in both cases, a fully compositional Content is available, but is ignored here):

50. $[_{N1} [_{N2} \sqrt{\text{DRAGON}}] \sqrt{\text{FLY}}]$

			CONTENT AND COMPOSITION:
i.	en-search 1	$[_{N1} [_{N2} \sqrt{\text{DRAGON}}] \sqrt{\text{FLY}}]$	no listed Content $[\sqrt{\text{FLY}}]$
ii.	en-search 2	$[_{N1} [_{N2} \sqrt{\text{DRAGON}}] \sqrt{\text{FLY}}]$	no Content match FLY
iii.	Unified en-search:	$[_{N1} [_{N2} \sqrt{\text{DRAGON}}] \sqrt{\text{FLY}}]$	DRAGON-FLY

51. $[_{N1} [_{N2} \sqrt{\text{WALK}}] \sqrt{\text{MAN}}]$

			CONTENT AND COMPOSITION:
i.	en-search 1	$[_{N1} [_{N2} \sqrt{\text{WALK}}] \sqrt{\text{MAN}}]$	no Content match $[\sqrt{\text{MAN}}]$
ii.	en-search 2	$[_{N1} [_{N2} \sqrt{\text{WALK}}] \sqrt{\text{MAN}}]$	no Content match no Content match
iii.	Unified en-search:	$[_{N1} [_{N2} \sqrt{\text{WALK}}] \sqrt{\text{MAN}}]$	WALKMAN

Because Content was matched with $[\sqrt{\text{FLY}}]$ in (50), the compound is 'endocentric', where by *endocentric* here we mean that the compound is a sub-kind of its head (i.e., a *dragonfly* is a *fly*). $[\sqrt{\text{DRAGON}}]$, however, receives no independent Content here and hence cannot serve as a Contentful restriction on *FLY*, except insofar as it delimits the en-search for what would turn out to be a *DRAGONFLY*. As a consequence, a unified en-search would return a unique atomic Content for *dragonfly*, *DRAGONFLY*, just in case it is consistent with the Content already matched previously with *FLY*, a situation I encode through the hyphen in (50), i.e. *DRAGON-FLY*. In contrast in (51) no Content was ever matched with the head $[\sqrt{\text{MAN}}]$, nor, for that matter, with the non-head $[\sqrt{\text{WALK}}]$. As a result, the compound is 'exocentric' where by *exocentric* here we refer to the fact that *walkman* is not a sub-kind of *man*. Crucially, the variability of Content matching in all these cases and most importantly the emergence of 'partial' compositionality, as in *dragonfly*, is made possible by the optionality of Content matching, together with the view of C-core as a domain in which at most one Content matching is possible.²³

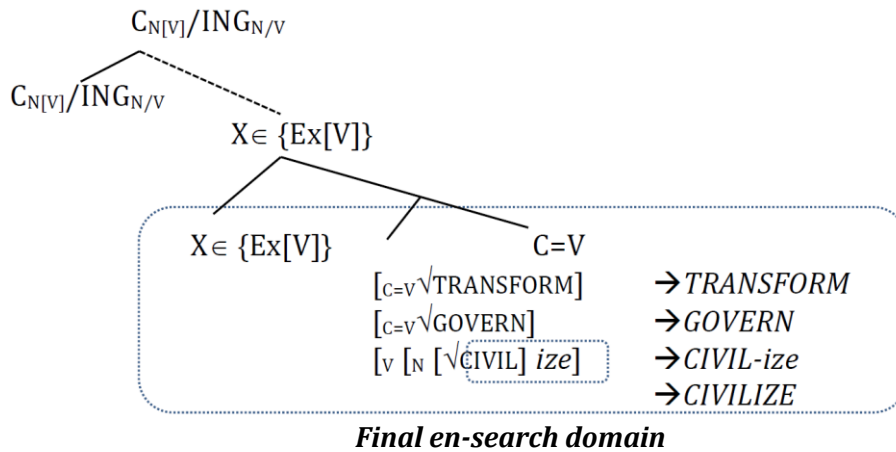
²³ 'Phrasal' idioms clearly cannot be subsumed under this system. Not only do they contain, potentially, functional material, some idiomatic expressions consist of nothing but, e.g. *be it as it may*. In turn, there are excellent reasons to reject the claim that non-compositionality for 'words' and non-compositionality for 'phrases' should be handled by the same formal mechanisms, including, but not limited, to the discontinuous nature of phrasal idioms, to the availability of anaphoric binding and arguably movement within phrasal idioms, and most importantly, to the fact that it is not at all clear that

6 Puzzles 1 and 2: Resolution

6.1 AS-nominals must be Compositional

Consider again the structure of AS-nominals, as in (37), repeated here, essentially, as (52). The upper, final domain available for en-searching is the edge defined by the outer box in (52), at the merger of the lowest ExP-segment (and see fn. 20 on the inclusion of the lowest ExP-segment head). It thus emerges that Content must be matched with some qualifying constituent within that box, or the relevant constituent is relegated to permanent Content limbo. Within that domain, $\sqrt{\text{TRANSFORM}}$ and $\sqrt{\text{GOVERN}}$, or for that matter *civilize* have not yet merged with $C_{N[V]}$. They may be matched with Content, compositional or not, as such, but clearly, non-compositional, atomic, Content cannot be matched with *transformation*, *government* or *civilization*, as such Content matching would require a non-local en-search. Atomic, non-compositional Content can, of course, be matched with *civilize*, with either the Content *CIVIL-ize*, or with the Content *CIVILIZE*, as in (52). Beyond that, however, the merger of *civilize* with a nominalizer must give rise to compositional Content:

52. AS Nominals:

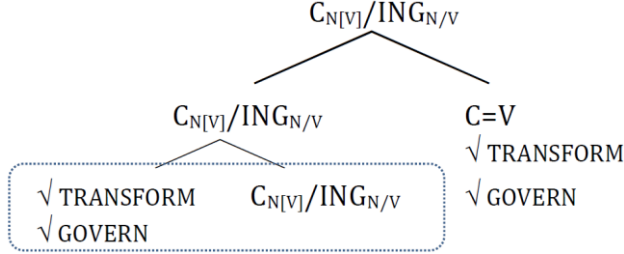


most phrasal idioms are, indeed, fully non-compositional in the relevant sense. Some pivotal examples of such effects are in (i-v) (all examples from Nunberg, Sag and Wasow 1994)

- i. *Partial compositionality, anaphoric binding:*
 - a. We_i will cross that bridge_j {when/if and when/at such time as} we_i {come/get} to it_j
 - b. {throw out/toss out/damp/junk/discard} the baby with the bathwater
 - c. {ramming/shoving} something down somebody's throat
- ii. *Internal embedding:*
 - a. leave no legal stone unturned; tilting at the federal windmills;
 - b. birds of a feather have a strong preference to flock together
- iii. *Elided parts of phrasal idioms:*
 - a. My goose is cooked but yours isn't ____
 - b. We thought *the bottom would fall out of* the housing market, but it didn't ____
- iv. *Coordination and object sharing with non-idiomatic verbs:*
Reinventing and tilting at the Federal windmill
- v. Phrasal idioms that are *obligatorily* the output of movement rules (passive, tough, comparatives):
 - a. *The die is cast; The race is run; If the truth be known; My goose is cooked; fit to be tied; caught short; written on water; when all is said and done; cast in stone; may as well be hung for a sheep as for a lamb*
 - b. *hard to take; play hard to get; too hot to handle; difficult to come by*
 - c. *more dead than alive; colder than a witch's tit*

No such restrictions hold for R-nominals. Here, a single C-core can contain both the stem and the C-functor without an intervening ExP-segment, allowing a single en-search to successfully return atomic Content for both *government* and *transformation*:

53. *R-nominals*:



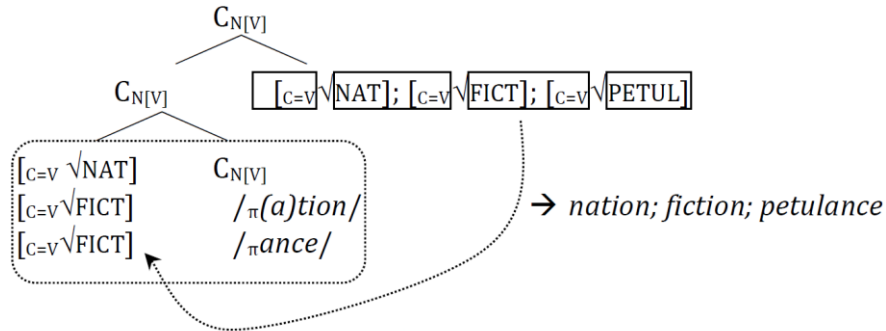
6.2 Puzzle 2, a Fresh Perspective

Puzzle 2, recall, consisted of the observation that AS-nominals may not be constituted of stems that are not, themselves, 'real' phonologically faithful verbs ((42b) repeated here as (54)):

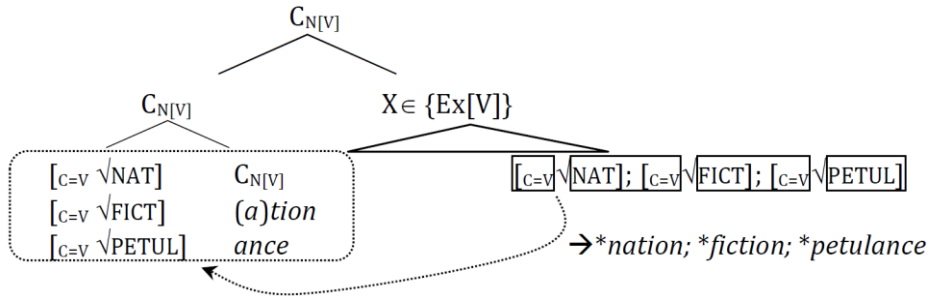
54. {Ex[V]} (or some subset of it) enforces a possible 'stand-alone' phonologically faithful V. Why? (Puzzle 2)?

The generalization, on the face of it, is puzzling. Specifically, and repeating the structures in (40)-(41), although the extended projection {Ex[V]} in the AS-nominal in (56) certainly defines its complement as V-equivalent (C=V), so does the C-functor $C_{N[V]}$ in (55). Why, then, must such a C=V constituent be a potential 'stand-alone' in (56) but not necessarily in (55)?

55.



56. *



The puzzle, as stated, becomes even trickier when compounded with what now emerges as the residue of puzzle 1. In section 6.1. I derived the obligatory compositionality of AS-nominals from the fact that Content matching cannot be accomplished past the merger of the first ExP-segment. It certainly emerges, as a result, that non-compositionality is excluded for AS-nominals. What does not emerge, however, is that they need to have Content altogether. Specifically, if Content matching is across-the-board optional, then why must the embedded V,

however categorized, have Content altogether, or differently put, why do AS-nominals need to have Content?

Or do they? Let us assume, rather, that as already suggested, Content matching is always optional, and that as a consequence, it is optional in (56) as well. What emerges is that if the C=V constituent in (56) is not matched with Content, the AS-nominal becomes a nonce form rather on a par with *slithy*. That this indeed may be the case can be illustrated by using the very same roots used, e.g. in (34), but this time providing them with event structure and arguments:²⁴

57. a. Kim's *moring* of the chocolate for five minutes; the *moring* of the chocolate (by Kim) in order to tempt the students
- b. Mary's *gribance/morification* of the chocolate with caramelized candy in order to counteract its bitterness

It thus emerges that while AS-nominals certainly must be compositional, if the relevant C=V constituent is not matched with Content, such compositionality is effectively vacuous. What crucially remains of it is the generalization that regardless of whether Content is matched with the C=V constituent, an AS-nominal may not acquire a Content independently of it – if the C=V constituent is a nonce form, so must the AS-nominal be.

But if the C=V constituent within AS-nominals may be devoid of Content, thereby allowing *the morification of chocolate*, then puzzle 2 as stated in (54) becomes rather incoherent. Certainly, *morify* and *gribe* are not 'real' or 'stand-alone' in any coherent sense that could possibly distinguish them from *fict* or *nat*?

Still, if the generalization in (54) is to be abandoned, how can we rule out the cases in (58) which gave rise to the generalization to begin with (cf. (15) and related discussion)?

58. a. *the vision of the mountain by the instructors for three hours this morning
- b. *the cognition about the problem for three hours by the philosophy student
- c. *the petulance of the passenger for thirty minutes in order to gain access to the lounge
- d. *the fiction of the book in order to win a prize

A direct answer to the conundrum is available once we consider the contrast between the cases in (58) and the cases in (59), the latter based on the same C=V constituents as those in (58). The cases in (59), we note, are licit AS-nominals. Rather than ungrammatical, they are nonce AS-nominals, on a par with the cases in (57):

59. a. the *nating/natance/natement* of the kittens by the neighbor
- b. the *ficting/fictance/fictal* of the book (by the student) in order to win a prize
- c. the *vising/visment/visal* of the mountains by the traveler

We may now assume, as seems clearly warranted, that en-searching returns no Content for [_C√NAT] or [_C√FICT], and that as a result, any AS-nominal derived from them is similarly Contentless. What, then, is the difference between e.g. (59b) and (58d) such that the latter, but not the former, is to be ruled out?

An answer is available, as it turns out, when we consider the fact that *fiction* can be matched with atomic Content, if an R-nominal, but not so *fictance* or *fictal*. Specifically, within R-nominals, the picture of Content matching to e.g. *fiction* and *fictal* is as follows:

²⁴ Note that if AS-nominals, complete with event arguments, are possible in the context of nonce verbs, then it emerges directly that the obligatoriness of 'stand-alone' verbs within such nominals cannot be derived from verbal listing plus inheritance, along the lines outlined in section 2.2.

- | | | | | |
|-----|-----|---------------------|----------------------------------|---|
| 60. | i. | $[_{C\sqrt{FICT}}]$ | | no listed Content |
| | ii. | a. | $[_{N[C=\sqrt{FICT}]} C_{N[V]}]$ | \rightarrow <i>fiction</i> <i>FICTION</i> |
| | | b. | $[_{N[C=\sqrt{FICT}]} C_{N[V]}]$ | \rightarrow <i>fictance</i> no listed Content |

We note, before proceeding, the crucial role played, in the matching of Content in (60), by the phonological realization of the C-functor, a matter I return to in section 7. Focusing on the contrast between (58) and (59), note that the cases in (58) are only ungrammatical if e.g. *fiction* and *vision* are associated with the atomic Content *FICTION* and *NATION* respectively. Such atomic Content matching is excluded in AS-nominals, as already noted, thereby rendering the presence of event structure in (58) illicit. It therefore follows that to be able to match Content with *fiction* and *vision*, it is necessary to construe their occurrence as coming with the syntactic structure of R-nominals. If, however, the structure is that of R-nominals, the occurrence of the subsequent event properties, be they arguments or event modification, is ungrammatical. We are faced here, in other words, with a classical *garden path* effect. To ‘make sense’ *fiction* and *vision* must be R-nominals, but under that parse, (58a-d) are ungrammatical. To become grammatical, in turn, would require overlooking the listed Content of *nation*, *vision* etc., so as to replace it with the equivalents of *the act of ficting* or *the act of vising*, which, on the whole, are hardly informative utterances. Under the assumption that preferring grammatical, but Contentless strings over Contentful ones, albeit ungrammaticality is at the very least difficult, if not impossible, the semblance of ungrammaticality emerges.

No such problems emerge when encountering *fictance* or *natement* and similar. Here, no Content can be matched with *fictance* or *natement*, or, for that matter, with *fict* or *nat* under any derivation. The derivation is thus free to proceed along any grammatically possible lines, be they those of R-nominals or AS-nominals, without any conflict between Content and structure. The result, sure enough, is a nonce form and hence not particularly informative, but this, as we already saw, is not a grammatical impediment in and of itself, Content matching being altogether optional.

Puzzle 1 and puzzle 2 do, then, reduce to the same factor: because AS-nominals involve event structure, and because in AS-nominals the stem and the nominalizer come together across ExP-segments, AS-nominals must be compositional (relative to the Content of the verb) or forever be doomed to Contentlessness. Insofar as Contentful outputs are favored, if such Contentful outputs are incompatible with AS-nominal structure, the appearance of ungrammaticality emerges.

Importantly, and most crucially, what emerges is that the interaction between the domain of en-searching, and hence Content matching, and complex words must be sensitive to the internal syntactic composition of words, and two strings which are morpho-phonologically identical may nonetheless place distinct constraints on en-searching because of the distinct syntactic structure that is embedded within them. Inevitably, then, complex words *must* be syntactically complex constituents, and cannot be otherwise derived or explained.²⁵

²⁵ While this article presents, I believe, an extremely strong argument for the syntactic derivation of derived nominals, any attempt to return derived nominals to the syntax cannot be complete without tackling the numerous arguments for a lexical treatment mounted originally in Chomsky (1970). Such a full account is not attempted here quite simply for reasons of space, but see Borer (2013b), where these are discussed in great detail.

7 The Phonological Domain of Content: Puzzle 3

Consider now again the matching of Content in (60), and specifically, the fact that atomic Content can be matched with $[N[C=V\sqrt{FICT}] C_{N[V]}]$ when realized as *fiction*, but not with $[N[C=V\sqrt{FICT}] C_{N[V]}]$ when realized as *fictance*. The point was noted, already, relative to contrasts such as those in (37a). Specifically, consider the cases in (61):

61. a. the *government/governance* of the people by the people (truth equivalent)
- b. the *government/governance* was incompetent (truth equivalent or divergent)

(61) as well as the parallel case in (60) are at the core of what was referred to, in the conclusion to section 4, as puzzle 3. Specifically, we note that in (61a), an AS-nominal, the use of *government* or *governance* gives rise to utterances which are truth equivalent. Such truth equivalence emerges, presumably, from the fact that AS-nominals must be compositional, and thus the Content of both *government* and *governance* must emerge from composing with the Content matched with $[C=V\sqrt{GOVERN}]$. It therefore emerges from the truth equivalence of *government* and *governance* in AS-nominals that the distinction between *-ance* and *-ment* is in and of itself structurally and Content-wise insignificant, and rather, *-ance* and *-ment* are but distinct phonological realizations of the same C-functor, $C_{N[V]}$.

The picture is by no means unique to AS-nominals, and R-nominals as well may allow for compositionality, under which *government* and *governance* remain synonymous, yet again suggesting that *-ance* and *-ment* are neither in themselves differently Contented or structurally distinct. However, as an R-nominal, and exclusively so, *government*, but not *governance* may have the non-compositional Content *GOVERNMENT*. Our puzzle 3, then, can be formulated as follows: if *-ance* and *-ment* are no more than alternative Contentless spellouts for $C_{N[V]}$ in the context of $[\sqrt{GOVERN}]$, how can *GOVERNMENT* emerge in the context of *-ment* but not in the context of *-ance*? Differently put, how could the structure $[N[C=V\sqrt{GOVERN}] C_{N[V]}]$ give rise to the Content *GOVERNMENT* when pronounced *government* but not when pronounced *governance*?

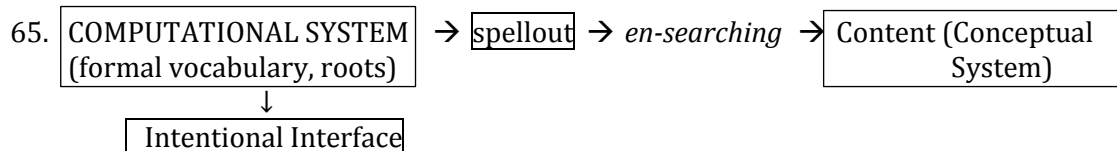
There is nothing special, we note, about *government* and *governance*, nor are other cases restricted to occurring vs. nonce forms such as Contented *fiction* vs. nonce *fictment*. Thus consider the following cases. In each case, the AS-nominal allows for more than one spellout for $C_{N[V]}$, with truth-equivalent variations emerging. In each case, however, only *one* such spellout (underlined) allows for non-compositional Content:²⁶

62. a. the selective *transmission* of historical documents
- b. scanning and *transmittal* of documents or parts of documents
- c. a camera system for processing documents for measurement of reflectance and/or *transmittance* of documents
- d. *transmission* → *GEARBOX*
transmittance → ***
transmittal → ***
63. a. Several groups ... monitor the sale and *transportation* of seed
- b. The *transportal* of seeds in the wool or fur of quadrupeds.
- c. *transportation* → (public) *TRANSPORTATION*
transportal → ***

²⁶ Because of rather strong idiolectal preferences for affixal realization, examples are based on corpora search (Google) where possible, and represent a best faith attempt at as minimal pairs as possible.

64. a. the slight *transference* of red pigments from the skins²⁷
 b. *transferral* of bread “sponge” from dough mixer to trough prior to fermentation
 c. *tránsference* → *TRANSFERENCE*
transferral → ***

Viewing all such cases, the picture that emerges is inevitable, as much as it is somewhat startling. Clearly, the Content-neutral phonological realization of C-functors *matters* when it comes to Content matching. Specifically, en-searching must be able to distinguish between two structures which are syntactically identical purely on the basis of the phonological realization of what is otherwise an identical syntactic constituent. The input to en-searching, then, constitutes not only the phonological index of the root, but also the phonological realization of C-functors, by assumption an instance of late insertion (Vocabulary Insertion, within Distributed Morphology). There is, simply, no other way to model the interaction between the spelling out of functors and the differently emerging Content.²⁸ If we continue to hold now, as is standard, that formal semantics specifically is neither sensitive to phonological realization nor triggers it in any way, the inevitable conclusion is that meaning, broadly construed, is not a unified component in Natural Language. While formal computations continue to take as their input syntactic representations devoid of phonological information, conceptual Content is en-searched specifically on the basis of phonological representations. The schematic picture that emerges, is thus as in (65):



The picture in (65), finally, amounts to a particular interpretation of Frege's (1892) distinction between *Sinn* and *Bedeutung*, with *Sinn*, as already noted, corresponding to Content, and *Bedeutung* to formal semantics – whatever meaning emerges from the computation of

²⁷While occurrences of the verb *transfer* in English with psychological Content, do occur, as well as, marginally, *transferral*, these are summarily rejected, as it turns out, by psychoanalysts (**transferral and counter-transferral*), who further insist specifically on an initial stress pattern (*tránsference*) as linked to the relevant term, although Standard English allows for *transférence* as well.

²⁸At least in principle, one could argue that it is the structure $[[_{C=V} \sqrt{\text{GOVERN}}] C_{N[V]}]$ that may be matched with the Content *GOVERNMENT*, and that the spellout *government* emerges as the result of already matched Content, rather than the other way around. That this is not the case, however, emerges from a consideration of cases, rather uncommon in English but quite frequent in e.g. Hebrew, where a structurally identical string consisting of $[[_{C=V} \sqrt{\text{XYZ}}] C_{N[V]}]$ gives rise to more than a single case of atomic Content, and where each of these instances of Content results in a distinct spellout, such as in (i):

i. $[_{N[C=V\sqrt{\text{NAT}}]} C_{N[V]}] \rightarrow \begin{matrix} \text{NATURE} \\ \text{NATION} \end{matrix} \rightarrow \begin{matrix} \text{nature} \\ \text{nation} \end{matrix}$

While the execution in (i) is certainly feasible, we note that it requires postulating ambiguous Content for $[_{N[C=V\sqrt{\text{NAT}}]} C_{N[V]}]$, together with a specific listing which matches e.g. *NATURE* with *nature* and *NATION* with *nation*. As such, it hardly by-passes the crucial connection between spellout and Content. It is also considerably less deterministic and more complex. By contrast, the matching of Content with spellout, while certainly requiring a list matching Content with spellout, doesn't need to postulate an ambiguity for $[_{N[C=V\sqrt{\text{NAT}}]} C_{N[V]}]$. For a fuller discussion, as well as for the picture in Hebrew, where the Contenting of spellout (rather than the spelling out of Content) appears inevitable, see Borer (2013b). Among other matters, I also propose a formal explanation for the relative rarity of pairs such as *NATURE/NATION* in English, vs. their endemic availability in Semitic. That very same account also results in accounting for the statistical – although not formal – favoring of the root domain in English for Content matching.

rigidly designating semantic functors, as well as whatever aspect of meaning emerges from syntactic architecture (e.g. scope) as mediated, by assumption, by C-functors.

8 Summary

Taking as our starting point the properties of AS-nominals and R-nominals headed, potentially, by morpho-phonologically identical forms, I noted two rather puzzling distinctions between these nominal instantiations which cannot be directly reduced to the fact that the former denotes an event while the latter need not do so. These puzzles, specifically, concerned the fact that AS-nominals, but not R-nominals must be compositional, and that derived nominals with a stem that is not, itself, an independently attested verb cannot be AS-nominals.

As it turned out, both puzzles can be resolved rather naturally and without much formal cost, but only if we assume that the internal syntax of AS-nominals and R-nominals is distinct. The relevant distinction does not just involve the structural properties of the phrase within which the nominal is embedded, but rather, involves the very internal structure of the 'head' itself. In other words the expression *transformation*, morpho-phonologically identical in R-nominals and in AS-nominals, nonetheless emerges as a result of a very different structure and as a result of very distinct derivational processes. More specifically, while both structures involve the merger of the verb (or verb-equivalent) constituent *transfer* with a functor to spell out as *-ation*, in the case of R-nominals, that merger is direct. In the case of AS-nominals, on the other hand, the verb and the nominalizing affix are separated by a verbal extended projection which, in turn, licenses both event interpretation and event arguments. The inevitable conclusion, then, is that the internal structure of words cannot be relegated to a component which is distinct from the syntax.

The return of word formation to the syntactic component requires, quite independently of the subject matter of this article, an explication of how non-compositional Content is matched with syntactic constituents, all the more so as compositional Content or its absence were at the core of the puzzles presented here. To that end, this article adopts the model in Borer (2009a,b, 2013a,b) according to which there exists a local syntactic domain within which non-compositional Content is available, specifically through an interface mechanism of en-searching (encyclopedic searching) which matches well-defined grammatical representations with conceptual Content. More specifically, I proposed that the relevant syntactic domain is the C-core. Differently put, within any extended projection, non-compositional Content is available to that portion of the structure which is embedded under the 'functional sequence', the set of nodes I have referred to here as ExP-segments (for segments of extended Projections). Much of this article was thus specifically devoted to establishing the validity of that domain, as based on the distribution of Contented items as well as the distribution of Contentless items, and the conditions under which both could be manipulated. What has emerged as a result of that discussion is that the availability of event structure and event arguments for AS-nominals and their obligatory compositionality follow from one and the same thing: the presence of syntactic structure which licenses both event interpretation and arguments, and which is instantiated, specifically, as ExP-segments which are part of the extended verbal projection. Such nodes certainly do license event structure, but at the same time, they also delimit Content to the C-core of the verbal extended projection, barring non-compositional Content for the combination of the verb and the nominalizer.

It thus emerges that the difference between (de-verbal) AS-nominals and R-nominals does not involve the presence vs. absence of a verbal constituent, clearly possible in both, but rather, from the presence in the former, but not in the latter, of segments of the verbal extended projection. The consequences of this result, we note, go well beyond the properties of derived nominals to suggest that roots, or verbs for that matter, are not, in and of themselves, responsible for the emergence of event arguments and event structure, but rather, properties of

event structure, and including the availability (or lack thereof) of both internal and external arguments derive from the presence of functional, grammatical structure. It emerges, as well, that lexical listing cannot possibly be the source of grammatical information concerning argument structure.

Finally, once the appropriate domain of Content was articulated and described, it emerged, rather inevitably, that the input to en-searching is not only syntactically delimited, but also phonologically determined. Specifically, it turned out that the phonological realization of functors which are otherwise devoid of Content may nonetheless impact the emergence of Content, thereby giving us a special meaning for *government* (i.e. *GOVERNMENT*) and *transmission* (*TRANSMISSION*), but not for *governance* (only *GOVERN-ance*) or *transmittal* (*TRANSMIT-al*). If on the right track, this suggests, in turn, that meaning, in natural language, is not a unified component. Rather, formal meaning, semantics, proceeds from the rigid computation of functors and structures, while Content is matched on the basis of phonological realization within a well-defined syntactic domain.

References

- Arad, Maya. 2003. "Locality constraints on the interpretation of roots". *Natural Language and Linguistic Theory* 21:737-78
- Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.
- Bobaljik, Jonathan. 2011. "Universals in Comparative Morphology", MS. University of Connecticut.
- Borer, Hagit. 1999. "The Form, the Forming and the Formation of Nominals." paper presented at the 2nd Mediterranean Morphology Meeting, September 1999.
- Borer, Hagit. 2003a "The Grammar Machine." in Artemis Alexiadou, Elena Anagnostopoulou and Martin Everaert. Eds. *The Unaccusative Puzzle*, Oxford University Press.
- Borer, Hagit. 2003b. "Exo-skeletal vs. Endo-skeletal Explanations: Syntactic Projections and the Lexicon." Maria Polinsky and Jonathan Moore. Eds.. *The Nature of Explanation*. Chicago: Chicago University Press. distributed by CSLI..
- Borer, Hagit. 2005a *In Name Only: Structuring Sense Vol. I. ..* Oxford: Oxford University Press.
- Borer, Hagit. 2005b *The Normal Course of Events: Structuring Sense Vol. II.* Oxford: Oxford University Press.
- Borer, Hagit. 2009a. "Compounds: Afro-Asiatic, Semitic: Hebrew." in R. Lieber and P. Štekauer. Eds. *Handbook of Compounds*. London: Basil Blackwell. 491-511.
- Borer, Hagit. 2009b "Roots and categories," Extended handout of paper presented at the 19th Colloquium on Generative Grammar; April 1-3, 2009, University of the Basque Country.
- Borer, Hagit. 2013a. "The syntactic domain of Content." In *Festschrift for Nina Hyams*, Amsterdam: John Benjamins.
- Borer, Hagit. 2013b. *Taking Form: Structuring Sense Vol. III*, Oxford: Oxford University Press.
- Carroll, Lewis, 1871 *Through the Looking-Glass, and What Alice Found There*
- Chomsky, Noam. 1970. "Remarks on nominalization." In R.A. Jacobs and P.S. Rosenbaum. Eds. *Readings in English Transformational Grammar*. 184-221. Waltham, Mass: Ginn
- Embick, David. 2003. "Locality, Listedness, and Morphological Information." *Studia Linguistica* 57:3, 143-169.
- Embick, David. 2010. *Localism vs. Globalism in Morphology and Phonology*. Cambridge, Mass: MIT Press
- Embick, David, and Morris Halle. 2005. "On the status of stems in morphological theory." In T. Geerts & H. Jacobs. Eds. *Proceedings of Going Romance 2005*. Amsterdam: John Benjamins. 59-88
- Frege, Gottlob. 1892. "On Sense and Reference" in A.W. Moore 1993. Ed. *Meaning and Reference*. Oxford: Oxford University Press
- Gajewski, Jon. 2010. "L-Triviality and Grammar." Talk presented at the University of Southern California, April 2010.
- Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, Mass: MIT Press.

- Halle, Morris and Alec Marantz. 1993. "Distributed Morphology and the pieces of inflection." In Hale, Kenneth and S. Jay Keyser. Eds. 111-176.
- Harley, Heidi. 2009. "The morphology of nominalizations and the syntax of vP." DP, QP and Nominalizations. Monika Rathert and A. Giannakidou. Eds. Oxford: Oxford University Press.
- Harley, Heidi. To appear "On the identity of roots". Theoretical Linguistics.
- Kripke, Saul. 1980. Naming and Necessity. Cambridge: Harvard University Press
- Levin, Beth and Malka Rappaport Hovav. 2005. Argument Realization, Research Surveys in Linguistics Series, Cambridge University Press, Cambridge, UK.
- Marantz, Alec. 1997. "No escape from Syntax: Don't try morphological analysis in the privacy of your own lexicon." In Alexis Dimitriadis et al. Eds. UPenn Working Papers in Linguistics 4.2.: 201-225
- Marantz, Alec. 2000. "Roots: the universality of root and pattern morphology." Paper presented at the Conference on Afro-Asiatic Languages, University of Paris VII, June 2000.
- Marantz, Alec. 2001. "Words." Keynote address to the West Coast Conference on Formal Linguistics, University of Southern California, February 2001.
- Newmeyer, Fritz. 2009 "Current challenges to the lexicalist hypothesis: an overview and a critique." In Will Lewis, Simin Karimi, Heidi Harley, and Scott Farrar. Eds. Time and Again: Theoretical perspectives on formal linguistics In honor of D. Terence Langendoen. Amsterdam: John Benjamins.
- Nunberg, Geoffrey, Ivan Sag and Thomas Wasow. 1994. "Idioms." Language 70: 3, 491-593
- Rappaport-Hovav, Malka and Beth Levin. 1998. "Building verb meanings". In M. Butt and W. Geuder, eds., The Projection of Arguments: Lexical and Compositional Factors, CSLI Publications, Stanford, CA, 97-134.
- Roeper, Thomas. 1987b. "Implicit Arguments and the Head-Complement Relation", Linguistic Inquiry 18:267-310.
- Veselinova, Ljuba, N. 2006. Suppletion in Verb Paradigms: Bits and Pieces of the Puzzle. Amsterdam: John Benjamins.
- Williams, Edwin. (1987). "The theta structure of derived nouns", in Need, B., E. Schiller and A. Bosch (eds.) Proceedings of the Chicago Linguistic Society, 366-76

© Hagit Borer
September 2013
London, UK

h.borer@qmul.ac.uk